Remote sensing-based approaches for large-scale comprehensive assessments of tree cover and windbreaks in the Great Plains region of the United States

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Dacia Marie Meneguzzo

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Dr. Joseph F. Knight, Advisor

August 2020

Acknowledgements

I especially want to thank to my advisor, Dr. Joseph Knight, and my committee members, Drs. Mark Nelson, Mark Hansen, and Dean Current for their help and not giving up on me during this long journey. Thanks to my co-worker and friend, Greg Liknes, for all the help, advice, and support.

Dedication

This dissertation is dedicated to my husband, Josh. You are a true partner in life and your support in everything I do means the world to me.

Abstract

Trees are an important resource in the Great Plains region of the United States yet little information describing their extent and location is readily available in formats that are convenient for resource professionals and decision makers. National forest inventory and natural resource monitoring programs seldom account for these non-traditional forests in their official statistics. In addition, most satellite-derived datasets are too coarse to accurately depict small or narrow groupings of trees common in the Great Plains. As a result, there is a lack of scale-appropriate data for inventory and monitoring of these tree resources.

Methods are needed to conduct large-scale comprehensive assessments of tree cover in the Great Plains. Remote sensing-based approaches offer several advantages over ground based inventories because they are often cost effective, they alleviate access issues, and they provide wall-to-wall spatial coverage.

The research presented here will demonstrate that tree cover can be mapped at a statewide level using an object-based image analysis (OBIA) approach and high-resolution (i.e., 1 m) digital aerial photography from the National Agriculture Imagery Program (NAIP) as the sole data source. Initial results indicated that the OBIA method was more accurate in terms of describing the actual observed spatial pattern of tree cover and produced a more realistic output product compared to a pixel-based classification method. Next, technological improvements were made to the OBIA method to make it more robust for operational land cover mapping at a regional level. Lastly, a shape-based classification approach was developed for positively identifying various configurations of windbreaks (both single and multiple-leg) from the output land cover maps, which is an improvement over existing methods that only map single-leg windbreaks. This is important for management purposes since windbreaks provide many ecological and economic benefits on the landscape,

from conserving topsoil to protecting crops, livestock, and farmsteads from the harsh effects of wind.

The outcomes of this research are actual published (or in the process of) high-resolution geospatial data products that are publicly available for download. These datasets identify and provide detailed spatial information about mapped tree cover and windbreaks that can be summarized at a variety of scales, from individual farms to the state or regional level. In addition, they are valuable for many different types of research studies and on-the-ground management activities. In a region of climate extremes, the hope is that these datasets will support informed decision making for placing trees in the right place on the landscape to maximize the benefits they can provide. For example, one of the goals in this region is windbreak establishment in areas with highly erodible soils that lack trees arranged as windbreaks. These maps will assist with such planting efforts as stated by Darci Paull, a GIS technician with Kansas Forest Service, “If we know where windbreaks are, then we know where they aren’t. Combining this information with other spatial information, for example, highly erodible soils data, we can identify at-risk soils that would benefit from the protection of a windbreak.”

Table of Contents

List of Tables

List of Figures

Introduction

Management of forest resources is dependent on forest inventory data. Such data is often obtained through national forest inventory (NFI) efforts. Treed resources in the agriculturally-dominant Great Plains region, however, are not like traditional forests. They may be remnants of former forests or intentional plantings but they have one thing in common: their location, or placement, is often intended to provide a specific ecological service, such as conserving soil, protecting crops, livestock and humans, or enhancing water quality. They occur in various formations from single, scattered trees on pasture and range lands, to linear plantings along fields, roads, or around farmsteads, to sinuous corridors alongside streams and rivers, or in small woodlots. While these tree features often occur on lands that fail to meet formal definitions of “forest” used by national inventory programs, their importance is recognized by the region’s land owners and resource managers. Windbreaks are a prime example; they are a critically important resource yet little information describing their extent and location is available in formats (e.g., digital maps) that are useful for natural resource professionals and decision-makers.

According to the *Assessing the Sustainability of Agricultural and Urban Forests in the United States* report, agricultural forest data are lacking, which undermines the ability to quantify ecosystem services. This is especially true for windbreaks and riparian forest buffers, which are the most commonly established agroforestry plantings because they provide important benefits that have a positive impact beyond the land they are planted on (Robertson and Mason 2016). The report also states that the lack of data for these tree resources is because they are not explicitly inventoried by either of the two primary natural resource inventories that fall under two agencies within the United States Department of Agriculture (USDA): the Forest Inventory and Analysis (FIA) program of the U.S. Forest Service and the Natural Resources Inventory (NRI) conducted by the Natural Resources Conservation Service (NRCS). Anecdotal evidence also suggests that windbreaks are being removed at a high rate but

there are no recent data to support or contradict these claims, especially for the multi-state Great Plains region. Notes from the Great Plains Windbreak Initiative Workshop held in Manhattan, Kansas, on February 7-9, 2017, state that, “There is a critical need for expanded, coordinated, region-wide efforts to raise the visibility of and address the need to monitor, renovate/manage, and establish windbreaks.” However, this cannot be done without the necessary and appropriate data products. Natural resource professionals and decision-makers in these heavily managed landscapes need reliable and scale-appropriate data for management and conservation tree planting efforts.

There are two primary ways in which data can be obtained about the Great Plains tree resource: measuring trees on the ground or using a remote sensing-based approach to identify and extract tree cover from satellite or aerial imagery. Sample-based approaches that collect data on the ground only provide summarized information, such as total area or length. Furthermore, the only spatial information provided is a point location, which does not provide sufficient information for conducting meaningful spatial analyses. Regardless, additional multi-state ground-based data collection is not an option due to the high expense, but a remote sensing-based approach offers a more timely and cost-effective solution that produces wall-to-wall spatial information. Such an approach also alleviates issues with obtaining landowner permission and accessing remote areas; however, there are drawbacks with existing land cover products. For example, the National Land Cover Database (NLCD) Tree Canopy Cover (TCC) datasets are obtained from 30-m spatial resolution satellite sensors that are too coarse to resolve these “sub-forest” areas that include single tree crowns, linear rows of trees (e.g., windbreaks) that are less than 30-meters wide, or patches of trees that are smaller than 900 m^2.

Publicly available digital aerial photography from the National Agriculture Imagery Program (NAIP), on the other hand, offers a potential solution and has several advantages for this particular type of application: the spatial resolution (≤1 m) is high enough that so that individual trees are easily identifiable, it is

available for the entire U.S. on a state-by-state repeat basis every 2-3 years, and it has spectral bands capable of separating tree cover from other features. However, working with this imagery is very challenging because of processing large data collections with variable quality amongst the many images, which is caused by changing atmospheric and illumination conditions encountered during the several month-long acquisition process. Furthermore, illumination conditions can vary even during the same day as sun angle changes, e.g., between two adjacent flight paths. This results in spectral differences in the final image products, which can negatively impact classification results. While the challenges of working with NAIP imagery are well documented in the literature, few solutions are offered for using NAIP imagery to conduct comprehensive assessments of tree cover across the large expanse of agricultural and rangelands in the Great Plains region of the United States.

This research was inspired by the desire to help staff at the USDA National Agroforestry Center (NAC) in Lincoln, Nebraska, answer three questions they are often asked about tree resources in the Great Plains: how much tree cover is out there? Where is it? What is it doing? The overall objective of this dissertation is to develop operational remote sensing-based approaches to answer these questions and improve our understanding of tree resources in the agroecosystems of the Great Plains region. This dissertation is organized into 5 chapters: a literature review, three studies in their own chapters, and a conclusions chapter. The objectives of chapters 2 and 3 are aimed at answering the first two questions while the focus of chapter 4 is to begin to answer the what-is-the-tree-cover-doing question. The three studies presented in Chapters 2-4 are manuscripts that have been or will be published in a scientific journal.

Chapter 1 is a literature review that provides the historical background of national forest inventory and the challenges of inventorying “trees outside forests” from the U.S. Forest Service’s FIA program’s perspective. The history of land cover mapping is also discussed as well as how the new era of high spatial resolution data and object-based image analysis (OBIA) procedures can be used

to address many land cover classification topics, such as mapping tree cover in agricultural landscapes.

Chapter 2: *Mapping trees outside forests using high-resolution aerial imagery: a comparison of pixel- and object-based classification approaches* (Meneguzzo et al. 2013) is a peer reviewed paper published in Environmental Monitoring and Assessment; it is licensed for inclusion in this dissertation. The objective of this chapter was to investigate a pixel-based approach (independent component analysis) and an object-based image analysis (OBIA) approach as potential solutions for broad-scale mapping of all tree cover in agricultural landscapes from very high-resolution imagery, i.e., 1 m NAIP imagery. We concluded that while both were viable approaches to mapping tree cover over broad spatial extents, OBIA was the better choice because it produced classification results that were more accurate in terms of spatial location that also provided more reasonable and realistic information about the spatial pattern of tree cover. This publication has been cited 58 times according to Google Scholar.

Chapter 3: *Developing statewide high-resolution land cover maps for assessing tree resources in the central United States using NAIP imagery (intended outlet is Remote Sensing)*. The objective of this research is to develop a robust and transferable method for operational mapping of general categories of land cover in rural landscapes using 1 m NAIP imagery as the sole data source for a four-state study area (Kansas, Nebraska, North Dakota, and South Dakota) in the Great Plains region of the central United States. An additional goal was to develop a single segmentation routine that can be applied within the four-state study area without having to make adjustments when moving from county to county or state to state. This chapter expands upon the OBIA method developed in Chapter 2. Technological improvements are implemented to make the approach more automated and transferable to other geographic areas. The segmentation routine was improved upon and we achieved the goal of developing one segmentation routine that was applied in all four states in the study area without making any adjustments to the segmentation parameters.

This method was applied in an operational manner to all four states in the study area and resulted in the publication of the first ever 1-m land cover datasets for Kansas and Nebraska; maps for North Dakota and South Dakota are forthcoming. The Kansas and Nebraska datasets are available for download on the USDA Forest Service Research Data Archive (Kansas: https://www.fs.usda.gov/rds/archive/catalog/RDS-2017-0025) (Nebraska: https://www.fs.usda.gov/rds/archive/catalog/RDS-2019-0038). While most OBIA land cover mapping studies are limited in geographic extent, i.e., less than 3 km^2 (Ma et al. 2017), this study concludes that areas larger than 200,000 km^2 can be mapped using an OBIA approach and NAIP imagery as the sole data source.

Chapter 4: *Identification of windbreaks in high-resolution land cover maps from the Great Plains region of the USA using a shape-based classification model (intended outlet is International Journal of Remote Sensing)*. The objective of this study was to develop a practical, operational method for identifying both single and multiple-leg treed windbreaks in high-resolution (1 m) land cover maps for large geographic regions. In particular, we used the high resolution land cover dataset produced for Kansas in Chapter 3 to attribute tree cover into a functional category, i.e., windbreaks, based on shape characteristics. Shapes metrics are robust in that they are not affected by image quality issues and I hypothesized that they can be used to positively identify windbreaks across landscapes with varying densities and patterns of tree cover. While most windbreak mapping studies consider only single-leg windbreaks (i.e., simple, regular linear features), this study includes multi-leg windbreaks as well, which have been problematic when it comes to windbreak classification (e.g., Liknes et al. (2017) and Ha et al. (2019)). The other limitation in windbreak mapping studies is similar to the issue of small study areas as mentioned above in OBIA land cover mapping; most windbreak mapping studies are also limited in geographic extent. For example, Pasher et al. (2016) and Vannier and Hubert-Moy (2014) both note that semi- or fully automated windbreak detection/mapping methods have not been implemented in an operational manner for large geographic extents. The

windbreak classification method developed in this chapter is an incremental improvement in identifying multiple-leg windbreaks and achieves the goal of developing a semi-automated approach that can be applied over large geographic areas, e.g., more than 200,000 km^2. The result is the creation of the first high-resolution, comprehensive statewide 1-m windbreak maps for Kansas. This dataset provides spatial information that describes each individual windbreak as well as composite information that describes the collective windbreak resource, such as total area and length of windbreaks for an entire state. Results for Kansas show that there are more than 50,000 km of windbreaks occupying nearly 106,000 hectares throughout the state. This endeavor is the first of its kind in the region and provides information at a new scale that is appropriate for inventory, monitoring, and decision-making regarding windbreaks and the tree resource as a whole.

Chapter 5 summarizes conclusions from the literature review (Chapter 1) and three publishable studies (Chapters 2, 3, and 4) that focus on developing remote sensing-based approaches for conducting comprehensive assessments of tree cover and windbreaks in the agriculturally-dominant Great Plains region of the United States. This region is a vast one, spanning an area nearly 500 to 1,000 km wide that extends 4,800 km in length from Canada to Texas and includes parts of ten states (Montana, North Dakota, South Dakota, Wyoming, Nebraska, Kansas, Colorado, Oklahoma, Texas, and New Mexico) and three provinces (Manitoba, Saskatchewan, and Alberta) (Robinson and Dietz 2019). The work presented in this dissertation is constrained to a subset of this region, specifically the states of North Dakota, South Dakota, Nebraska, and Kansas. The developed methods may have utility for conducting assessments of tree cover in other agriculturally-dominant regions but have not been tested beyond the Great Plains.

Chapter 1. Literature Review

1.1 Introduction

Tree cover in the central United States today consists primarily of trees established for agroforestry purposes and those naturally occurring along stream and river corridors. Federal and other tree-planting movements led to the establishment of tens of thousands of miles of windbreaks and shelterbelts throughout the Great Plains during the early- to mid-20th century (Hurt 1996). Although scarce in terms of overall coverage, these tree resources provide a wide variety of important functions (Guo et al. 2004), such as: protection of soil, livestock, crops, and wildlife; increase crop productivity; reduction in energy inputs and chemicals; improvement of water quality; providing more economic opportunities; and adding to the biodiversity and structural diversity of the landscape (Rietveld and Irwin 1996). More recently, these small tree fragments have also been recognized for their significant role in greenhouse gas mitigation by sequestering carbon (Schoeneberger 2009; Kort and Turnock 1999). However, due to their small size and/or narrow shape, these “trees outside forests” (TOF) are excluded from broad-scale forest inventory programs that employ a formal definition of forest land. So, despite their importance, little information about TOF is currently available and national inventory and monitoring programs are lacking (Kleinn 2000; de Foresta et al. 2013). The overall goal of this study is to develop a comprehensive and broad-scale remote sensing-based assessment of tree cover and windbreaks in the Great Plains region of the United States.

1.2 National Forest Inventory and Trees Outside Forests (TOF)

The configuration of TOF in agricultural landscapes presents a challenge to traditional inventory procedures since it often occurs in long, narrow bands or small, disconnected patches. National forest inventory (NFI) programs are responsible for obtaining information about forest land over large areas and

consequently rely on sample-based approaches in order to be cost effective. Circular plot-based sampling that is typical of NFI's is not ideal for inventorying tree cover in areas where linear configurations dominate; line intersect sampling is more appropriate for these types of features (Hansen 1985). Furthermore, NFI's often utilize a formal definition of forest land that determines which tree-covered lands will be included in the inventory. In the United States, the Forest Inventory and Analysis (FIA) program of the USDA Forest Service defines forest to be land with a minimum of 10% tree cover (or equivalent stocking) that is at least 1 acre in size and at least 120 feet in width (U.S. Forest Service 2010). Consequently, the configuration of tree resources in agricultural landscapes often precludes them from inclusion in the inventory. For example, Perry et al. (2009) concluded that including non-definitional forest land with trees in the inventory would increase estimates of total tree-covered land by at least 25% in Kansas, Nebraska (26%), North Dakota (38%), and South Dakota (30%).

Interest in TOF is not new. For example, the 1955 periodic inventory report for Nebraska by Stone and Bagley (1961) reported on the area of wooded strips and windbreaks and also included stand-size class information for these trees. The next inventory report (1983) by Raile (1986) provided more detailed information pertaining to TOF, such as area by nonforest tree class, forest type, site index, basal area, stand age, and even ownership, while the report for the last periodic inventory in 1994 (Schmidt and Wardle 1998) provided less detailed information. The aforementioned reports obtained this TOF-related information via statistical sampling procedures and aerial photography (M. Hansen, personal communication, April 17, 2013). Other examples of information collected on TOF in the United States include Hartong and Moessner (1956) in Iowa and Hansen (1985) in Kansas but these studies only considered wooded strips. After the establishment of the new annual forest inventory system via the 1998 Farm Bill, the first phase of the FIA inventory included collecting information about "nonforest treed lands" by aerial photo interpretation personnel. This information was used in the previously mentioned study by Perry et al. (2009) and various

FIA inventory reports to report on the area of nonforest treed lands that are not included in the FIA inventory (e.g., Meneguzzo et al. 2008). However, these estimates are still sample based and comprehensive coverage and spatial information is still missing from these efforts.

The most recent large-scale effort was the Great Plains Tree and Forest Invasives Initiative (Great Plains Initiative; GPI) program implemented in North Dakota, South Dakota, Nebraska, and Kansas in 2008 and 2009 to account for trees that are not included in the FIA inventory in order to address the potential impact of invasive pests such as the emerald ash borer (Lister et al. 2012). Using a sample-based approach, it was estimated that there are 5.1 million acres of nonforest lands with trees compared to 6.4 million acres of forest land, indicating trees occupy similar areas of non-definitional forest land and forest land in the four-state region. As part of the GPI inventory process, the functions of the nonforest trees were recorded and estimates of total area by function were determined. For example, there are more than 500,000 acres of field windbreaks in the four states (Meneguzzo et al. 2018). While this brings us closer to accounting for trees outside of definitional forest land, the GPI field inventory was a one-time effort and there is no spatially descriptive information associated with these estimates, and the associated sampling errors can be very high: 100% or higher in some cases.

As we look over the history of the FIA inventory program, it becomes clear that it has evolved from a timber-based inventory to one that encompasses all forest land (albeit definitional forest land), and now we are beginning to see the interest shift to a more comprehensive inventory that includes all tree-covered lands, which was suggested by Perry et al. (2009). In addition, Strategic Goal 3 of the *Integration of the USDA Agroforestry Strategic Framework, Fiscal Year 2011-2016* calls for “Work within USDA to establish a comprehensive continuous national inventory of on-the-ground applications of agroforestry practices/systems or include in existing inventory structures (e.g., Forest Inventory and Analysis or the National Resources Inventory).” However, as

previously mentioned, most agroforestry plantings are excluded from the FIA inventory due to their size and/or configuration. As such, new inventory procedures will be required to even begin to obtain such information.

1.3 Definition of TOF

"The term Trees outside forests, a neologism coined in 1995, is framed in the forest context, defining the concept by default with reference to forested areas. So an exact definition of the term requires a reading of the definition of the term forest, which, as we know, varies from country to country in accordance with the environmental stakes, economic interests and local situations involved." (Bellefontaine et al. 2002)

There is no globally-accepted, formal definition of TOF. The Food and Agriculture Organization of the United Nations (FAO), who has pioneered the effort to inventory TOF on a global scale, uses the following definition: "trees on land not defined as forest and other wooded land" (FAO 2001); examples include trees that occur on agricultural and grazed lands, along waterbodies and roads, and in residential and urban settings (Rawat et al. 2003). So, in their global assessment of TOF, it would include trees that do not meet their formal definitions of forest or "other wooded land". For detailed definitions, see de Foresta et al. (2013). In the United States, there is no formal definition of TOF either. By process of elimination, FIA is the only national forest inventory in the United States that includes all land ownerships, so it would be logical to refer to all trees that do not meet FIA's definition of forest land as TOF.

1.4 Characteristics of TOF

TOF differs from traditional forest land in that TOF can be found anywhere trees can grow (de Foresta et al. 2013). When we think of traditional forest land, we tend to think of larger contiguous tracts of tree-covered lands in rural areas. However, this is not the case with TOF. They are found in urban as well as rural areas and may consist of one single tree or a group of trees; and, they may occur

naturally or have been planted in a specific configuration in order to serve a specific purpose, such as windbreak. In terms of spatial pattern, they can be scattered or found systematically around farmsteads and agricultural fields (Kleinn 2000; de Foresta et al. 2013). Another distinct characteristic of TOF is that they are often established to serve a specific ecological and/or economic function. As such, they have been referred to as "working trees" (Perry et al. 2009) and the National Agroforestry Center has published a series of *"Working Trees"* papers to aid landowners in correctly establishing trees to achieve a particular goal, such as enhanced water quality or improved crop yields. Overall, there is a lot of variation amongst TOF in terms of extent, configuration, spatial pattern, and function.

1.5 Importance of TOF

Despite their occurrence and importance in urban settings, this study will focus on tree cover in rural areas only. Rural TOF provide a multitude of important ecological and economic functions. As mentioned earlier, they offer protection to soil, livestock, crops, and wildlife; increase crop productivity; reduction in energy inputs and chemicals; improvement of water quality; providing more economic opportunities; and adding to the biodiversity and structural diversity of the landscape (Rietveld and Irwin 1996). Other uses listed by de Foresta et al. (2013) include: timber, fuelwood, spatial markers of ownership boundaries, living fences, fruits, nuts, shade, aesthetics, and odor control.

Perhaps one of the more timely issues though is carbon sequestration. While forests are often considered and studied as large carbon stores, TOF are gaining recognition as important carbon stores as well (Brandle et al.1992; Kort and Turnock 1999; Schoenebeger 2009). In fact, The Carbon Benefits Project: Modeling, Measurement, and Monitoring effort developed guidelines for "measuring carbon stocks in above ground woody biomass in non-forest land

cover." (Castaneda et al. 2011, p.6). Furthermore, the process they developed relies less on field data and more heavily on satellite remote sensing.

1.6 Assessing TOF

In order to properly manage TOF, decision- and policy-makers need to know the extent, location, function, and contribution of these resources. However, broad-scale inventory and monitoring programs to do such assessments are lacking (Kleinn 2000; de Foresta et al. 2013). For example, only one study is currently available that partially assesses TOF on a global scale: Zomer et al. (2009) found that "more than 10 million km^2 (46 percent of the land classified as agriculture land in the global datasets) have more than 10 percent tree cover" (de Foresta et al. 2013, p. 63). While the results of this study are somewhat "rough estimates" given the coarse resolution of the datasets used in the study, the results still support the key idea that TOF on agricultural lands are an important feature globally (de Foresta et al. 2013). While only a handful of countries, e.g., India, directly incorporate TOF into their forest inventories, this notion is gaining momentum and FAO continues to be the global leader in pushing the effort to inventory and assess TOF at all scales, especially broad scales.

1.6.1 TOF assessment methods

While no TOF assessment procedures have been officially implemented, de Foresta et al. (2013, p. 56) identify three primary methods that are used alone or in conjunction to provide information about TOF:

1. "Remote sensing and the analysis of aerial photographs and satellite images (combined with ground checking, may provide information on the extent, localization and spatial organization of TOF. Impressive technological progress has made remote sensing as essential tool for measuring these parameters and their change with time.")

2. “Field inventories that combine sample plots with various tree measurements for information on the tree resource itself; species, etc. Provided the sampling scheme is adapted to the area covered, valid statistic estimates of the tree resource over the whole area can be derived, such as the number of trees, the stocking volume, the carbon stock, etc. “
3. “Survey questionnaires may provide information on various aspects of the TOF resource especially on land used for tree crops in agriculture, but also on urban land.” Information obtained is some limited.

They also state that land use/land cover assessments are a more commonly used and important source of information for a TOF, or any natural resources, assessment. Furthermore, they stress that it is important to think “operationally and incrementally.” Klein (2000) also identifies aerial photos and satellite imagery as suitable methods for land use classification and mapping for acquiring information about TOF. Similarly, Johansen (2007, p. 30) points out with regards to mapping and assessing riparian and neighboring forest ecosystems using high-resolution satellite imagery that, “Spatially extensive and non-invasive remote sensing techniques are therefore often applied due to their synoptic and repetitive nature and ability to be utilized in areas which are not easily accessible.”

While TOF can be found anywhere trees grow, FAO divides them into TOF on: 1) agricultural land, 2) settlement land, and 3) land not used for either purpose mentioned in 1 and 2. Furthermore, they acknowledge that assessments of urban and rural lands are conducted using their own specific and separate methodologies (de Foresta et al. 2013). In accordance with this idea, this study will focus on TOF on agricultural land, or rural areas in general. This is supported by recent work by Meneguzzo et al. (2013) who found that when using an image-based land-cover assessment approach to map TOF, urban areas were very complex and required their own classification model.

Selecting an approach for assessing TOF in the central U.S. has several important considerations: rural areas are quite extensive, TOF are scattered across this large area, and current budgets do not allow for additional expensive field data collection. Given these considerations and the three options presented above, Option #1 is the most applicable.

1.7 History of remote sensing and land cover mapping

Remote sensing, aerial photography in particular, has a long history in forest inventory and land cover mapping applications. According to Loveland (2012), “The historical roots of land-cover mapping reside in the early history of aerial photography and applications spanning forestry, agriculture, urban planning, and water-resources management.” And Franklin (2001, p. 18) writes, “The technology of remote sensing originated in the science and technology of aerial photointerpretation (Silva, 1978).”

The beginning of aerial photography interpretation is largely associated with World War I because this was when aerial photography acquisition began on a regular basis (Campbell 1996). Research related to the nonvisible portion of the electromagnetic spectrum took off during World War II, and remote sensing research and applications in the civilian sector continued after the War (Campbell 1996). One important example is work by Colwell (1956), which is considered “Among the most significant developments in the civilian sphere” and “forms a clear milestone in the development of the field of remote sensing” because it “applied color infrared aerial photography, first used widely during World War II, to important problems in the plant sciences” (Campbell 1996, p. 7).

The 1940s and 1950s witnessed the development of radar and thermal imagery, respectively, and in the 1960s, meteorological satellite data became available on a more routine basis (Franklin 2001); it was from the meteorological satellite that land observing satellites were later developed (Campbell 1996). Another significant development in the 1960s was the coining of the phrase

“remote sensing” by Evelyn Pruit, a scientist for the U.S. Navy’s Office of Naval Research (Campbell 1996).

While the 1960s witnessed rapid growth in satellite technology, many developments were also made in aerial photo interpretation and photogrammetry. In the early part of the decade, it was established that object recognition in an image could be accomplished using six basic properties: size, shape, shadow, tone and color (reflectance), texture and pattern (Rabben 1960; Colwell 1965), and, by 1965, automated methods had been developed to extract these six ‘interpretation elements’ from an image. While it was originally thought that only skilled photo interpreters and photogrammetrists could accurately do this, Colwell (1965) dared to present an optimistic discussion about the possibility of machines extracting data similar to human interpretation. The conclusion was that human and machines were complementary for identifying objects and their context from aerial imagery.

In the post-war era, aerial photography was an important resource for mapping. According to Steiner (1965, p. 65), aerial photos, at and up until that point in time, were used “rather extensively as base maps for field mapping, very frequently in conjunction with land classification projects.” He provides several significant examples of land use mapping that utilized aerial photos (for a complete summary see Steiner 1965). The most significant work by far is that of Marschner (1958), who’s map titled “Major Land Uses in the United States, 1:5,000,000” was “the only example anywhere in the world in which an extremely small scale land use map for a whole country has been prepared directly from air photographs.” (Steiner 1965, p. 66). A similar example, although much smaller in extent, is the “Land Use Categories in Pennsylvania, 1:250,000” map (Klimm 1958), which mapped six land use classes. The California Vegetation Inventory was the result of an experiment that sought out whether vegetation mapping could be done from aerial photograph and Steiner (1965) describes this as the “earliest example in the United States of a vegetation and land use map compiled by actual photo interpretation.” Many of the other examples of early aerial photo

mapping consists of using the photo as a base map in the field (field mapping, often associated with natural resource surveys by government agencies), or flying in a plane and recording observations on previously acquired aerial photos (e.g., MacFadden 1949). Interestingly, by the time Steiner's paper was published in 1965, pre-war aerial photos were valued for their historical purposes and already being kept in the National Archives in Washington. Furthermore, land use change studies were already being conducted as early as 1941 (see Andrews and Bromley 1941).

In the late 1960s, aerial photo interpretation was believed to have earned the respect of researchers across a number of disciplines (Colwell 1968 *in* Franklin 2001), and then things changed drastically (Franklin 2001). The launch of Landsat-1 in 1972 and the availability of its new digital imagery changed the course of remote sensing research. New challenges were focused on making statistical manipulations of data (King 2002) and developing the new technology. It was the dawning of the 'digital remote sensing' era.

1.8 The "modern" era of remote sensing

As mentioned above, the early 1970s arbitrarily defines the "modern" era of land-cover mapping (Loveland 2012), or the "Landsat era" (Landgrebe 1997). While aerial photography had a strong presence in land cover and land use mapping, things changed with the launch of Landsat-1 in 1972. After this point, there was a great increase in land cover mapping using this new digital imagery and Loveland (2012, p. 14) points out that "land-cover studies are a key driver of the Landsat mission." According to Campbell (1996), the launch of Landsat-1 was another "milestone" and provided two important contributions to remote sensing research: first, multispectral data for large areas of the earth's surface were now available on a routine basis and this led to large increase in the number of scientists conducting multispectral analyses. Secondly, analyses of digital imagery were widely and rapidly expanded. Overall, the 1970s witnessed three prominent trends according to Loveland (2012, p. 16): "the development

and acceptance of computer-assisted land-cover mapping techniques, the growth of land-cover mapping initiatives across the United States and other parts of the world, and the improvement in Landsat data quality due to Thematic Mapper instrument."

Besides purely technical advances, perhaps one of the most legendary contributions to come out of this decade was the development of the standard land-use and land-cover classification system by Anderson et al. (1976), or 'Anderson's Classification System.' It consists of four levels so it can be applied at a local to national scale. In addition, it was designed so that the classes at each level would have per-class accuracy of at least 85% when mapped at the appropriate scale; this figure of 85% has become a land-cover mapping accuracy standard that is still utilized at the present time (Loveland 2012).

Landgrebe's "The Evolution of Landsat Data Analysis" (1997) provides an excellent summary of the origin and development of multispectral data analysis (synonymous with Landsat data) and how it spanned across disciplines to reach a wide variety of researchers and users. The following are highlights from this paper:

1. Original motivation/research need behind the development of Landsat was for agricultural purposes
2. Development of a scanner capable of collecting electronic data for 18 spectral bands was accomplished by the U of MI group, while the Purdue group developed the ability to put all sets of measurements from the scanner "into one vector for each pixel"
3. Analysis methods based on spectral characteristics rather than image characteristics were chosen for several reasons: economic considerations (higher spatial resolution was cost prohibitive), the large volumes of data that needed to be processed, and processing time, which led researchers to choose the digital approach
4. It was quickly determined that this digital remote sensing data could be useful for other areas of research besides agriculture. The monumental

1967 study, “Useful Applications of Earth-Oriented Satellites,” showed just how widespread interest in this type of data had become, and was a driving factor for developing “a satellite series to be known as Earth Resources Technology Satellites (and later as Landsat), the design of which began in earnest within the next year”

5. At this point, remote sensing research was combined and interrelated effort that involved engineering, science, and applications objectives. Figuring out how to collect and analyze data, understanding scene materials, developing operational uses, pattern analysis and feature selection algorithms, image registration, developing new instruments for lab and field work were all research endeavors, and led to the development of “several extensive databases of laboratory and in situ spectra;” another result of this research was the publication of serial journals and books dedicated to this new field
6. One of the earliest studies that used Landsat consisted of making county-level land use maps using the maximum-likelihood classifier for the Great Lakes drainage basin region. Other early significant examples of larger-scale efforts involving Landsat were LACIE (the Large Area Crop Inventory Experiment) and AgRISTARS (Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing). These studies had important contributions, such as “a means to label training samples for the supervised classifier in an unbiased fashion without observations from the ground, because these were not possible in this program.” Others included the Tasseled Cap transformation (Kauth and Thomas 1976) and promoted the use of temporal information in analyses.
7. The development of analysis methods comes from examining the data. Multispectral data have been represented and visualized in three fundamental ways: image space, spectral space, and feature space. Spectral space is associated with the commonly used term “spectral

> signature.” Approaches for mathematical representation of multispectral data include deterministic approaches (most common), stochastic models, fuzzy set theory, Dempster-Shafer Theory of Evidence, Robust Methods, Theory of Capacities, Internal Valued Probabilities, Chaos Theory and Fractal Geometry, A1 Techniques, and Neural Networks. Other important characteristics of analysis methods for Landsat data are identification versus discrimination (does the algorithm identify or distinguish between pixel information?), and determining how humans are involved in determining classes of interest for a particular analysis.

The 1980s saw significant improvements in land cover mapping and began what Loveland (2012, p. 17) calls “the large-area operational era for land-cover mapping.” During this decade, Landsat data quality improved with its new Thematic Mapper ™ sensor that provided more detailed data due to higher spatial resolution (30 meters) and increased multispectral capabilities, and there were a number of important advancements in multispectral classification, image processing, and software system development that enhanced processing in order to produce land-cover products. (Loveland 2012). Franklin (2001, p.18) also writes, “In 1985, Curran suggested that rapid technological advancements (e.g., computers) and improved sensor systems (both introduced and envisioned) had propelled remote sensing into a stage of exponential growth.” This is further evidenced by the appearance of statewide land-cover mapping programs (Loveland 2012).

In addition to Landsat, the French Satellite Pour l’Observation de la Terre (SPOT) was launched in 1986 (Campbell 1996) and aided land-cover mapping efforts with its 10- and 20-m high resolution data (Loveland 2012). While these sensors provided higher resolution data, coarse-scale resolution data acquired from the advanced very high resolution radiometer (AVHRR) was found to be useful on a much larger scale when it was used to map land cover for Africa

(Loveland 2012; but see Tucker et al. 1985). However, this was just a precursor to the flurry of remote sensing-related activity that occurred during the 1990s.

The study by Tucker et al. (1985) opened the door for a myriad of other national-scale land cover studies and mapping projects (e.g., Frederiksen and Lawesson 1992; Gaston et al. 1994; Cihlar et al. 1996; Zhu and Evan 1994; Loveland et al. 1995) as well as the development of the first global-scale land-cover products (e.g., DeFries et al. 1995; DeFries et al. 1998; Loveland et al. 1999) (Loveland 2012). These global-scale studies based on AVHRR data yielded new techniques, such as the use of the normalized difference vegetation index (NDVI) to add information for classification and alleviate the effects of using such coarse-resolution input data; utilizing seamless datasets to avoid boundary issues; improving classification accuracy by using ancillary data and stratification (Loveland 2012).

On a smaller scale, land cover mapping projects over large areas using high-resolution Landsat and SPOT imagery also grew (Loveland 2012). In the United States in particular, the formation of the Multi-Resolution Land Characterization (MRLC) consortium led to the expansion of national land-cover mapping. Perhaps the best known outcome was the USGS National Land Cover Database (NLCD) 1992 (Vogelmann et al. 1998), which was probably "the most widely used land-cover dataset in the United States" (http://www.epa.gov/mrlc/nlcd.html; accessed May 9, 2013). Successive datasets include NLCD 2001 (Homer et al. 2007) and NLCD 2006 (Fry et al. 2011). According to Wulder and Franklin (2007, p. 199), "The vast majority of published research involving landscape pattern analysis has used Landsat TM data. Reliable, repeat availability (since the 1970s) at low price (or free) has rendered it the true "workhorse" of landscape ecology." Important technical advances in classification included artificial neural networks (Hepner et al. 1990; Gopal and Woodcock 1994) and decision tree analysis (Friedl and Brodley 1997; DeFries et al. 1998; Hansen et al. 2000).

Throughout the 1990s, SPOT, Landsat, MODIS, and ASTER satellite systems were used in many important mapping and resource assessment projects at regional to global scales. However, the end of the decade welcomed the '1-m generation' of satellite sensors (IKONOS in 1999, QuickBird in 2001, and OrbView in 2003) that would make way for new satellite remote sensing applications that were previously only possible using aerial imagery (Blaschke 2010). This created a shift in the way we thought about image classification of such high-resolution data: from pixels to image objects (Blaschke and Strobl 2001). Along with the shift to object-based analyses, we see an increase in the use of Light Detection and Ranging (LiDAR) data in conjunction with high-resolution multispectral data.

The 2000s continued to see improvement in techniques for land-cover mapping over large areas and change detection efforts. Land cover products (500-m) were developed using MODIS data, including the Vegetation Continuous Fields product (Hansen et al. 2002). Other global land cover projects included Global Land Cover 2000 (GLC2000) and Globcover 2006, which is currently the "highest resolution global land-cover product currently available" (Loveland 2012, p. 19). Important NLCD-derived products such as impervious surface (%) and percent tree canopy also emerged (Homer et al. 2007). In terms of policy, an important decision was made in 2008 when the USGS Landsat Data Policy made all of their Landsat data freely available to the public via the internet. Suddenly, these data were readily available for a long time period, which will likely lead to new developments in multitemporal analyses (Loveland 2012).

1.9 Remote sensing and forest cover mapping

"During the past decade, digital remote sensing has become an increasingly important tool for mapping and monitoring forest resources around the globe." Cohen et al. (1996, p.421).

Remote sensing has played an important role in forest resource mapping. While aerial photos have a long history in forest inventory, monitoring, and management applications, Landsat data also came to play an integral role in such activities (Bauer et al. 1994). Cohen et al. (1996) provides a nice review of some of the significant research and examples of how digital remote sensing has been used in forest mapping applications in the Pacific Northwest region of the U.S. (Table 1.1).

Table 1.1 Summary of important studies related to the role of digital remote sensing in forest mapping applications. From Cohen et al. (1996).

Authors/year	Geographic Area	Data Sources	Objective	Accuracy
Walsh (1980)	Crater Lake National Park, OR	Landsat MSS	12 land cover types	88.8%
Isaacson et al. (1982)	Blue Mountains in OR	Landsat MSS; large-scale air photos	Elk habitat (vegetation type and some forest structural attributes)	Not reported
Cibula and Nyquist (1987)	Olympic National Park, WA	Landsat MSS: topographic and climate data	Map 21 vegetation cover classes	91.7%
Eby (1987) and later on Eby and Snyder (1990)	west side of the Cascade Range/western WA	Landsat MSS	map old-growth forests (age structure)	80% (Cascades of WA) 85% (Olympic Peninsula)
Morrison et al. (1991)	National Forests of the west side of OR and WA Cascade range	Air photos; Landsat MSS; Landsat TM; panchromatic SPOT; DEM data	map old-growth forests	Not reported

Congalton et al. (1993)	Similar to Morrison et al. (1991)	Landsat TM; Air photos; DEM data; field measurements	map old-growth forests	80-91% for nine National Forests
Fiorella and Ripple (1993)	Douglas fir forests	Landsat TM; topographic relief image from a DEM	map forest successional stages	78.3%
Ripple (1994)	10.9 million acres of forest in OR	AVHRR	Mapped percent confier; presented as an assessment of forest fragmentation	Not reported
Cohen et al. (1995)	1.24 million ha in western OR	Landsat TM	Mapped forest cover (4 classes of canopy coverage)	75% for three collapsed age classes; 82% for the full map

These examples clearly illustrate the important role of remote sensing in forest mapping and monitoring, and demonstrate how the use of satellite data has become operational in forest resource monitoring (Cohen et al. 1996; Franklin 2001). Bauer et al. (1994) also describe the operational use of Landsat data in the development of an Annual Forest Inventory System (AFIS) in Minnesota, USA. Many of these studies also incorporated the use of other datasets, such as aerial photos and digital elevation models (DEM), as well as data transformations and image processing techniques in order to improve classification accuracies. This is important because these developments went beyond using only spectral properties and contributed to the evolution of the field of remote sensing as it pertains to forest resource mapping and monitoring.

In more recent forest resource studies, we see the use of image segmentation with high resolution imagery as well as with Landast data for more

detailed analyses, mainly for mapping forest stands. For example, Dorren et al. (2003) compared pixel- and segmentation-based classifications of forest stand type maps in an area with steep terrain. This also illustrates the shift from pixel- to object-based approaches. Another example is Hay et al. (2005), who developed the multiscale object-specific segmentation (MOSS) approach for automatically segmenting high-resolution imagery into "meaningful forest-objects" at various scales, which range from single tree crowns to forest stands in a 1 km^2 study area.

It is not uncommon for high-resolution imagery to be used in support of large-area NFI efforts. It has been used in conjunction with field data (i.e., modeling procedures) to produce complete raster-based outputs of forest resources and small-area statistics in Sweden and Finland (Tomppo et al. 2008) and in support of updating existing forest inventory databases; Falkowski et al. (2009) developed a framework in which high-resolution imagery could be used for automated stand delineation and attribute estimation in Canada's large-area sample-based forest inventory, or adapted for use in other forest inventory and monitoring efforts at a national to global scale. In addition, the FIA inventory has always incorporated the use of high resolution imagery, from analog aerial photos in the past to digital imagery today, in the initial phase of the inventory. These are examples of the operational use of high-resolution remote sensing data in broad-scale forest inventory and monitoring efforts.

1.10 Remote sensing and land cover mapping to assess TOF in the United States

Given the need for an image-based TOF inventory outlined in Section 1.6, a broad-scale, operational land cover classification (with a focus on tree cover) will be utilized to obtain my research goal of a comprehensive assessment of tree cover in the central United States using the states of Nebraska and Kansas as the study areas.

1.10.1 Imagery and existing land cover dataset considerations

As previously mentioned, de Foresta et al. (2013) suggested existing land use/land cover datasets as means to assess TOF. While efforts have been made to use satellite imagery to comprehensively map land cover across the conterminous U.S. (e.g., the National Land Cover Dataset (NLCD 2006)), the sensors used are too coarse to discern small groups or narrow tree plantings and do not provide consistent estimates of total tree cover (Perry et al. 2009; Liknes et al. 2010). Likewise, Johansen et al. (2007, p. 30) state that "The application of satellite image data for mapping riparian vegetation with moderate spatial resolution image data from the Landsat and SPOT satellites has produced limited results, as the spatial resolution of the sensors is too coarse to delineate narrow bands of vegetation along streams (Congalton et al. 2002; Muller 1997). There are no other higher-resolution land-cover datasets available for the broad extent under consideration for this assessment.

Since it is necessary to be able to identify TOF on the imagery, it is clear that high-resolution (< 5m) imagery is required for this study; this is also suggested by Kleinn (2000). However, obtaining satellite data over large areas can be expensive and cost is an important consideration. As such, imagery from the National Agriculture Imagery Program (NAIP) is a much more affordable alternative. It is acquired during the growing season (leaf-on) primarily for agricultural compliance monitoring and has been captured on a routine basis since 2003. It meets the needs of the study because it available for the entire study area (e.g., nationwide) at a sufficient spatial resolution (i.e., 1-2 meters), has repeat coverage, and is available at no cost. Congalton (2010, p.456) calls it a "tremendous source of digital imagery that the analyst can readily digitally analyze." Single trees and small fragments of tree cover are readily identifiable from this type of imagery. It is interesting to note that when Bauer et al. (1994) conducted their study, they found "research by Coppin (1991) has indicated the computer analysis of multidate Landsat data (summarized above) offers a cost-effective alternative to reliance on aerial photography." Now, however, NAIP and

Landsat are both available for free, and NAIP offers a cost-effective alternative to satellite imagery that is comparable in spatial resolution.

1.10.1.2 Challenges in classifying high-resolution imagery

Traditional per-pixel classification methods have been found to be inadequate for dealing with complex high-resolution imagery even with the inclusion of texture measures and new methods for feature extraction are necessary (Blaschke and Strobl 2001; Culvenor 2003 *in* Chubey et al. 2006). The resolution presents a challenge since a higher spatial resolution leads to increased spectral variation of landscape features, which makes it more difficult to statistically separate classes using traditional pixel-based classification methods and thus reduces classification accuracy; this is known as the 'H-resolution problem' (Woodcock and Strahler 1987; Marceau et al. 1990 *in* Hay et al. 1996). A current solution is to use object-based image analysis (OBIA) procedures (Lang and Blaschke 2006), which consists of two primary components: image segmentation and classification (Blaschke et al. 2000; Definiens 2010, p. 10). Basically, OBIA is said to mimic human visual interpretation of an image, which results in the extraction of real-world objects as opposed to just classifying small individual pixels that often represent only a tiny part of landscape features.

1.10.2 History and development of object-based image analysis (OBIA)

While it seems to be relatively new, current OBIA has its foundation in many past remote sensing research studies. For example, earlier segmentation and classification models that used spectral and spatial information to create spatially similar clusters of pixels and then classified those rather than individual pixels include the BLOB classifier (Kauth and Richardson 1977), ECHO (Kettig and Landgrebe 1976; Landgrebe 1980), and AMOEBA (Bryant 1979). However, computing constraints prevented these types of classification techniques from becoming operational until recently. The use of present-day OBIA for extracting information from high-resolution imagery has increased markedly during the last

decade, primarily due to the release of eCognition software and having readily available high-resolution imagery (Blaschke 2010). The main difference between pixel- and object-based classification is that while per-pixel classifiers often utilize only spectral information, object-based classifiers use spatial information as well (Lillesand et al. 2008).

Early object-specific work by Hay et al. (1996; 2001; 2003; 2005) illustrated its usefulness in dealing with problematic issues often encountered with remote sensing imagery, such as increased spectral variability in high-resolution imagery and the modifiable areal unit problem (MAUP). For example, Hay et al. (1996) used an image-texture technique called the triangulated primitive neighborhood (TPN) method to describe the spatial interaction of trees in high-resolution imagery. This approach was a different kind of spatial filtering method than was commonly used at that time: it was object-specific as opposed to using a window of fixed size. The goal was to reduce the variance in the complex, high-resolution image and improve classification accuracy. Following this, work by Hay et al. (2001 and 2003) identified how image-objects help alleviate the MAUP and their usefulness in multiscale landscape analyses. Additionally, Hay et al. (2005) developed multiscale object-specific segmentation (MOSS), a tool to produce consistent segmentation results at a range of scales that was more cost-effective than manual delineation. Similarly, Burnett and Blaschke (2003) had developed multi scale segmentation/object related modeling (MSS/ORM) to create image objects at more than one scale simultaneously and then incorporate this information into the classification process. On the other hand, work by Lang and Langanke (2006) demonstrated that one level representation (OLR) is most likely adequate and easier to implement in certain cases; either way, the end goal of image segmentation is always to create ‘meaningful’ image objects (Blaschke 2010).

Perhaps one of the most influential OBIA researchers is Thomas Blaschke, from the Department for Geography and Geology, University of Salzburg, Salzburg Austria. Starting with their definitive paper entitled “What’s

wrong with pixels? Some recent developments interfacing remote sensing and GIS" published in 2001, Blaschke's work provided important foundational work that has continued to shape and influence present OBIA research, and it has been important to landscape ecology research as well. Much of Blaschke's work brings a very interesting perspective and new ideas to integrate remote sensing, GIS, and landscape ecology using OBIA, especially in dealing with the issue of scale and applying the multiscale concept to landscape-level analyses. Due to his initial OBIA-related work in the early 2000s, he is credited with starting this recent shift in thinking from pixel- to object-based approaches for classifying high resolution imagery and incorporating the idea of object-based approaches for dealing with the complexity of landscape-level systems, including structure and function and their interactions. His oft-cited work includes: Blaschke et al. 2000; Blaschke and Strobl 2001; Blaschke and Hay 2001; Blaschke et al. 2004; Blaschke and Lang 2006; Blaschke et al. 2008; and, Blaschke 2010. Other fundamental work that has formed the basis of present-day applied OBIA research and is much cited includes Baatz and Schäpe (2000) whose paper is about multiresolution segmentation and Benz et al. (2004) who present a thorough description of OBIA and its role in creating GIS-ready products. Through my own observations, however, I have noticed that in more recent OBIA research papers it appears that their work is not always cited and has become what I would call 'core concepts' of OBIA.

1.10.3 Advantages of OBIA

Because OBIA groups similar pixels into 'image objects' and uses these as processing units, this method offers several fundamental advantages over per-pixel approaches: 1) image objects can be created at various scales (e.g., from a single tree crown to groups of trees) (de Jong and van der Meer 2004; Hay et al. 2003), 2) the use of image objects alleviates the salt-and-pepper effect often encountered in pixel-based classifications (Yu et al. 2006; Blaschke et al. 2000), and 3) numerous attributes can be obtained from image objects, including

statistics such as mean and standard deviation using the DNs (Chubey et al. 2006), and as further explained by (Benz et al. 2004, p.240), “meaningful statistic and texture calculation, an increased uncorrelated feature space using shape (e.g., length, number of edges, etc.) and topological features (neighbor, super-object, etc.), and the close relation between real-world objects and image objects.” In addition, classification results based on image objects have been found to be more accurate than those from pixel-based procedures (Blaschke and Strobl 2001; Benz et al. 2004; Yu et al. 2006; Cleve et al. 2008; Platt et al. 2008; Myint et al. 2011; Perea Moreno and Meroño De Larriva 2012; Li and Shao 2013). Other important aspects of this technique are that it bridges a gap between information contained in the image and having a usable GIS product (Benz et al. 2004; Lang and Blaschke 2006; Yu et al. 2006; Lang et al. 2009; Blaschke 2010) and other data layers (raster and vector) can be included as additional information in the segmentation and/or classification process.

1.10.4 Image segmentation

Simply put, segmentation is the process that produces image segments, or objects (Benz et al. 2004). However, there is more to it than that. According to Congalton (2010, p.452), “The power of the segmentation process is twofold. First, the imagery is now divided into polygons that can, in many ways, mimic the polygons that may have been drawn by an analyst that was manually interpreting this same image. In this way, some of the additional elements of manual interpretation mentioned earlier in this paper become relevant for digital image analysis. Secondly, as previously mentioned, the creation of polygons results in a powerful addition of attributes about the polygons that can be used by the classification algorithm to label the polygons. Both these factors significantly add to our ability to create accurate thematic maps.”

While image segmentation is not new (Haralick and Shapiro 1985; Pal and Pal 1993; Ryherd and Woodcock 1996), its use in remote sensing and geospatial applications has only recently increased (Blaschke 2010; Yu et al. 2006). While

there are other segmentation algorithms available, this project employs the commonly used ‘multiresolution segmentation’ algorithm, which is a Fractal Net Evolution Approach (FNEA) embedded in the eCognition software (Baatz and Schäpe 2000; Hay et al. 2003; Laliberte et al. 2007; Myint et al. 2011) and is commonly used for “extracting land cover or man-made features from remote sensing imagery” (Definiens 2010, p. 18). Starting with one pixel, adjacent similar pixels are merged to form larger objects based on heterogeneity criteria (Yu et al. 2006; Benz et al. 2004). The formation of image objects is controlled by scale, color, and shape, and merging stops when the threshold set by the scale parameter would be exceeded but a larger scale parameter will result in larger image objects (Laliberte et al. 2007; Benz et al. 2004). Furthermore, multiresolution segmentation allows the user to extract features at various scales since image objects are formed by user-defined settings that minimize heterogeneity and are not based on the pixel size alone (Hay et al. 2003; Benz et al. 2004). This is an important advantage of OBIA because it allows for the creation of image objects at multiple scales, e.g., having objects that represent individual tree crowns as well as large patches of forest within the same image.

It has been stated that the scale parameter is the most important setting in the segmentation process because classification takes place on the image objects, (Benz et al. 2004; Tansey et al. 2009; Myint et al. 2011) and that good segmentation is essential for achieving accurate classification results (Su et al. 2008; Zhou and Wang 2006; Wiseman et al. 2009; Tansey et al. 2009). Good segmentation/selecting the appropriate scale parameter means finding the setting(s) that result in image segments that accurately represent the landscape features of interest (often referred to as “meaningful” image objects) and meet the study’s classification objectives, e.g., houses, roads, trees, etc. However, this is no small task, especially in an operational sense, and using a trial-and-error approach to find the appropriate segmentation settings can be very time consuming (Duro et al. 2012). Recent work by Dragut et al. (2010) developed a tool that is intended to aid the user in selecting an appropriate scale parameter

based on local variance, but this tool and other methods of scale parameter selection have not been fully implemented in popular image segmentation software, such as eCognition (Duro et al. 2012). So, at the time this research was being conducted, most studies continue to use prior experience and trial-and-error approaches to find segmentation parameters that are appropriate for their study (e.g., Chubey et al. 2006, Myint et al. 2011; Duro et al. 2012). In addition, studies often use more than one segmentation routine with various parameters (mainly scale) to create multiple levels of image objects. According to Duro et al. (2012, p. 263), “The use of image object information derived from multiple image segmentation scales has been shown elsewhere to produce better overall classification accuracies (Smith 2010), and better classification accuracies for individual land cover classes (Myint et al. 2011).”

1.10.5 Image Classification

Congalton (2010, p. 451) says “By far the greatest advance in classifying digital remotely sensed data in this century has been the widespread development and adoption of object-based image analysis (OBIA). Traditionally, all classifications were performed on a pixel basis. Given that a pixel is an arbitrary delineation of an area on the ground, any selected pixel may or may not be representative of the vegetation/land cover of that area.”

Texture has long been considered an important property for image classification (e.g., Haralick 1973). One of the primary differences between pixel- and object-based texture calculations is the window size used. In pixel-based approaches, a fixed-size square window is used yet in object-based approaches, the image objects themselves serve as the window within which texture is determined. Within-object texture is determined as well as “a higher order of texture that takes into consideration the spatial distribution of adjacent objects can also be generated” (Chen et al. 2012, p. 4441). Hay et al. (1996) recognized this advantage of using an object-oriented approach to calculating texture in order to reduce variance in high-resolution imagery of forest canopy for working

with traditional classifiers. This is one of the main advantages in present OBIA work.

As mentioned above, image segmentation is powerful because it creates polygons and the attributes of the polygons provide a great deal of information that can be used during classification. Liu et al. (2008) presents a more thorough description of this information; they identify three levels of features, or attributes, which can be obtained from the segmentation process and incorporated into the classification. They write (p. 462): “Level 1 features are properties of a single image segment, such as area, perimeter, shape, index, and a range of texture measurements (Herold et al. 2003), that can be considered to classify the segment. For example, Guo et al. (2007) used these features to more accurately distinguish between dead crowns and bare ground in an oak forest. Level 2 features focus on spatial relations between two objects, such as containment, proximity, an adjacency; these can also be used to classify an image segment or refine the classification result. For example, if an image segment that is spectrally similar to vegetation occurs next to an image segment that is classified to be a house, the first segment is likely to be a lawn. Level 3 features are spatial patterns in which more than two objects are involved and could be used to aid in classifying segmented objects.”

Overall, there is a wealth of information that can be used in the classification phase of OBIA approaches, and although it is not often listed as one of the primary steps in OBIA methods, Li and Shao (2013) list feature selection as one of the main steps in object-based classification for land use and land cover mapping. However, this can be a daunting task given the large number of features available for classification. Researchers use different methods to select object features for classification; some methods are subjective and are based on the investigator’s knowledge and prior experience with conducting object-based classifications in the area of interest (e.g., Laliberte et al. 2007; Duro et al. 2012), while others use decision trees or the random forests algorithm for feature selection (e.g., Yu et al. 2006). Duro et al. (2012, p. 270)

also note that, “In addition, faced with potentially hundreds of object features from which to select, the use of more advanced feature selection algorithms in object-based image analysis is gaining increasing attention (e.g., Yu et al. 2006; Chan & Paelinckx 2008).”

OBIA workflows are often “linear” in that they use segmentation to create image objects then use the image object attributes to carry out the classification (O’Neil-Dunne et al. 2011). Such workflows do not take advantage of the ability to incorporate contextual information in the classification process, which is a significant advantage of OBIA techniques (Hay and Castilla 2006); however, some studies have used limited context information, such as the relationship between super- and sub-objects (O’Neil-Dunne et al. 2011). Perhaps an impediment to using contextual information is that it is not readily available from the original input data, and therefore deviates from the linear approach we are used to with pixel-based approaches. It emerges from iterative processing which creates more and more contextual information that can be used to help classify the remaining image objects that are difficult to identify using more straightforward classification attributes, such as spectral and textural properties. Basically, incorporating contextual information attempts to mimic human recognition of features in an image as much as possible, but the process of actually doing this remains a challenge (O’Neil-Dunne et al. 2011).

OBIA classification can be conducted using a number of different classifiers (Duro et al. 2012). Duro et al. (2012) provides a good summary of the classification approaches used with OBIA, especially in studies that compared pixel- and object-based classifiers. K-NN is a commonly used classifier in OBIA approaches (e.g., Yan et al. 2006; Yu et al. 2006; Platt and Rapoza 2008; Myint et al. 2011; Dingle Robertson and King 2011) and maximum likelihood classifier has also been utilized (Platt and Rapoza 2008; Castillejo-González et al. 2009). The study by Duro et al. (2012, p. 260) used “relatively modern and robust supervised machine learning algorithms: decision trees, random forests, and support vector machines.” Classification and regression trees, or CART

classifiers have also been implemented with success in OBIA studies. These are discussed further in Section 1.11 (Classification tree analysis) of this paper. Other interesting classification approaches include Frohn (2006) who utilized landscape metrics in his land cover classification scheme, and O'Neil-Dunne et al. (2011) who used complex rulesets in eCognition to build and incorporate context information into the classification process. Object relationship modeling that utilizes expert knowledge and the spatial relationship among image objects has also been utilized as a classification technique (e.g., Blaschke et al. 2004); Liu et al. (2008) also utilized spatial relations between objects for mapping roads and moving vehicles in their classification process.

Within the eCognition software environment, there are two primary OBIA classification options: supervised fuzzy logic nearest neighbor (NN) and fuzzy membership functions (Walker and Blaschke 2008 *in* Perea Moreno and Meroño De Larriva 2012). The NN approach uses training samples and the "Nearest Neighbor" classifier, which "searches for the closest sample image object in the feature space of each image object. If an image object's closest sample object belongs to a certain class, the image object will be assigned to it." (Definiens 2010, p. 93). Fuzzy membership functions use "fuzzy rules" to evaluate an image object's membership to the different classes using a value between 0 and 1 (Benz et al. 2004), the higher the value in a specific class, the more likely it will be assigned to that class. Fuzzy rules are defined using "membership functions" in eCognition. Fuzzy rules allow more flexibility and image objects can be assigned to one class over another based on whether it's more "like" one class than another. Additionally, straightforward class assignment can also be carried out using thresholds to assign image objects to a particular class.

1.10.6 OBIA and mapping tree cover

A review of the literature reveals that high resolution imagery (≤ 5 m) and OBIA have been used in conjunction as an approach for mapping woody plant features in agricultural and other rural landscapes with success around the globe.

For example, Tansey et al. (2009) used OBIA to accurately identify hedgerows with a 2-meter minimum width from aerial imagery in Berkshire, UK. Aksoy et al. (2010) carried out a study in Germany, the Czech Republic, and Cyprus in which an object-based methodology was used to automate the process of identifying linear wooded strips in agricultural areas. Other related studies have been conducted by Davies et al. (2010) who used an OBIA approach to estimate juniper cover from NAIP imagery in Idaho, USA, Platt and Schoennagel (2009) whose study assessed changes in tree cover over a 60-year time period in Colorado, and Wiseman et al. (2009) where large shelterbelts were mapped over an area covering approximately 25,900 ha using very high-resolution (62.5 cm) aerial imagery in Manitoba, Canada. The authors concluded that an object-based method was very efficient for broad-scale inventorying of shelterbelts. However, these studies were confined to small areas while more recent examples that use OBIA techniques occur over much larger areas, such as 289,755 ha in North Dakota, USA (Liknes et al. 2010), 177,000 ha in California, USA for crop identification (Peña-Barragán et al. 2011), and 111,000 ha in Minnesota, USA (Meneguzzo et al. 2012). Furthermore, Liknes et al. (2010) and Meneguzzo et al. (2013) explored the use of high-resolution NAIP imagery and OBIA for mapping all tree cover rather than only focusing on definitional forest land. Both studies found that the FIA estimate of definitional forest land was much lower than the estimated area of all tree-covered lands, thus providing more evidence that tree resources are largely underestimated in the agricultural landscapes of the Midwest. The recent use of OBIA techniques offers a promising solution to the challenge of mapping fine-scale tree and other land cover features from digital aerial imagery in an operational manner over a large spatial extent in the central United States.

1.10.7 OBIA and mapping land cover

In terms of mapping land cover, OBIA has been used with success. For example, Cleve et al. (2008) mapped 4 classes of land cover (built area, surface

vegetation, tree/shrub, and shadow) in a wildland-urban interface (WUI) area that was approximately one square mile in size in Napa County, CA. They also compared the results from using a pixel-based classifier (ISODATA) and found that OBIA results were more accurate. Platt and Rapoza (2008) also found that OBIA produced classification results that were an improvement over those from a pixel-based classifier when mapping land use/land cover classes.

Many OBIA studies that produce a land cover classification are performed in urban areas. I think OBIA procedures are well suited for the high level of complexity found in imagery of urban landscapes. For example, Zhou et al. (2008) mapped five classes of land cover (buildings, pavement, coarse-textured vegetation (trees and shrubs), fine-textured vegetation (herbaceous vegetation and grasses), and base soil in Baltimore Metropolitan Area using OBIA and reported overall accuracies of 92.3% and 93.7% for images from 1999 and 2004, respectively. Moskal et al. (2011) also performed a land cover classification in part of Seattle, WA that included the following classes: buildings, grass, developed, impervious, shrub, tree, bare ground, water/vegetation, and other. They also included a number of data layers in the process: roads, building footprints, imagery (NAIP, QuickBird, NLCD, and oblique photography), and texture. They concluded that OBIA worked well for urban tree cover assessments. Myint et al. (2011) tested five different classifiers to map land cover in a part of Phoenix, AZ. They had seven land cover classes: buildings, unmanaged soil, grass, other impervious, pools, trees/shrubs, and lakes/ponds. They found that OBIA had high overall accuracy at 90.40% and worked much better than the pixel-based classifier (maximum likelihood). However, they also conclude that detailed urban land cover classification is difficult, especially over large images. Even though these land cover studies took place in urban areas, they still provide valuable ‘lessons learned’ and can contribute to the overall process of mapping land cover using high-resolution imagery and OBIA.

1.10.8 Other OBIA mapping studies

Other OBIA studies often focus on mapping one particular class or feature of interest. Some interesting examples include Johansen et al. (2009) whose goal was to map banana plantations, and Amorós López et al. (2011) who wanted to use OBIA to update the citrus GIS inventory of the Comunidad Valenciana region in Spain by classifying citrus parcels. Zhou and Wang (2006) mapped impervious surface in Rhode Island and Zhou et al. (2009) even used an OBIA approach to map areas of shade in an urban study area. Another emerging trend is the use of OBIA and LiDAR; for example, Hellesen and Matikainen (2013) used such a process to map tree/shrubs in an urban area and found that, in their case, mapping trees/shrubs was dependent on LiDAR. LiDAR is commonly used in studies that need structure (i.e., height) information, such as forest structure or biomass studies, identifying tree cover in urban areas, and it is particularly useful for providing high-resolution and accurate elevation data. While it may be useful for classification, LiDAR is not readily available everywhere. As such, I will not include it in this study. Furthermore, since I wish to extend this work over large geographic areas, I will keep the amount of input data to a manageable level and only use one input layer: NAIP imagery, which is already about 1 TB in size. Adding LiDAR data would greatly expand the volume of input data.

As far as OBIA studies that use NAIP imagery, there are very few, especially for large geographic areas. It has been utilized in some urban studies that mapped tree/vegetation cover (e.g., Moskal et al. 2011; Troy et al. 2012; Li and Shao 2013). Davies et al. (2010) and Maxwell et al. (2017) and (2019) provide good discussion of the issues and challenges of working with NAIP imagery. See Chapter 3, section 3.2 for an up-to-date listing of broad-scale land cover mapping studies, including those that use NAIP.

One aspect of the study by Li and Shao (2013) that makes it applicable to my research is that one of their goals was to use only one data source (i.e., 1-meter aerial imagery, and they decided to use NAIP imagery) for feature extraction.

They compared pixel- and object-based approaches to classify urban vegetation types. Their OBIA classification results were superior to the pixel-based results and they also state that (p. 785), “This study demonstrated an operational methodology for detailed vegetation delineation with a single data source.”

1.11 Classification tree analysis

One inherent aspect of present-day OBIA studies is the multitude of image object attributes that can be derived and subsequently used in classification. Some of the general categories of attributes include spectral, spatial/textural, geometric (e.g., shape), and contextual properties. Throughout the evolution of land cover mapping, we have seen how incorporating textural and other information such as vegetation indices improves classification accuracy over that obtained using only spectral values.

Images contain spatial information such as texture, context, pixel proximity, and geometric properties (Narumalanin et al. 1998 *in* Frohn 2006). While many of these aspects have been incorporated into land-cover classification research, little of this research has focused on utilizing shape (Frohn 2006). One of the advantages of OBIA is that shape as well as all of the aforementioned types of information are readily available and can be directly incorporated into the classification process as inputs into NN, for building membership functions, or used as thresholds for assigning image objects to a particular class. However, with numerous attributes available, the challenge lies in choosing which ones are most useful for distinguishing between land cover classes, especially those that are spectrally, or otherwise, similar.

Decision trees have been used with success in land cover mapping since the 1990s (e.g., Friedl and Brodley 1997). In the summary of a paper by Friedl and Brodley (1997), they highlight the important advantages of using decision tree classifiers for broad-scale land cover mapping, such as having the ability to identify important attributes for distinguishing between classes. In addition, they also say that they are efficient and robust, which makes them very useful “for

classifying the large volumes of data inherent in remote-sensing land cover-mapping problems" and making them suitable for operational use. There has also been an increase in the use of classification and/or regression trees in OBIA methods (Blaschke 2010); one of the reasons being image object attributes often fail a normality test so a nonparametric approach, making decision tree analysis an appropriate classification method.

Classification trees are effective tools for finding relationships between many predictor variables and a categorical response variable, such as a land cover class. Here are some important OBIA examples: Laliberte et al. (2007) used image object input for CART® to classify vegetation in arid rangelands and Yu et al. (2006) also used object-based input and CART for detailed vegetation classification. In a slightly different application, Chubey et al. (2006) used a series of decision trees to relate image object information derived from high resolution imagery to forest inventory parameters. Zhou and Troy (2008) also state that "regression tree algorithms" can be used to find threshold values for classification. More recently Laliberte et al. (2012) tested three different methods for feature selection to do detailed vegetation mapping (i.e., species level) and found the CTA was the best method for very high resolution data (sub-decimeter digital aerial imagery) and achieved relatively high classification accuracies.

1.12 Random Forests

As noted above, one of the advantages of OBIA is that a large number of attributes can be calculated for each image object using its spectral, textural, and other properties, as well as from incorporating other data sources. However, this presents a challenge for traditional land cover classification because the commonly used parametric classifiers are not appropriate for this type of complex, multisource classification (Richards and Jia 1999; Gislason et al. 2006); therefore, alternative classification methods are needed. Of the proposed ensemble classification methods, boosting (Freund and Schapire 1996; Schapire 1999) and bagging (Breiman 1994) are likely the most commonly used (Gislason

et al. 2006). The following discussion of these methods is from Gislason et al. (2006): While bagging trains many classifiers using boot-strapped samples from a set of training data, boosting uses iterative retraining and samples that were incorrectly classified are given higher weighting as the iterations continue. Although boosting is more accurate, it has disadvantages: it requires more time to run, may overtrain, and is sensitive to noise (Briem et al. 2002). So, during training, “the Random Forest algorithm creates multiple CART-like trees (Breiman et al. 1984), each trained on a bootstrapped sample of the original training data, and searches only across a randomly selected subset of the input variables to determine a split (for each tree node). For classification, each tree in the Random Forest casts a unit vote for the most popular class at input *x*. The output of the classifier is determined by a majority vote of the trees.” (Gislason et al. 2006, p. 295). Overall, in their assessment of RF compared to other ensemble classification methods for land cover classification, Gislason et al. (2006) conclude that RF achieves high accuracy, does not overfit the data, is efficient, and doesn’t need guidance from the user. It also estimates variable importance in the classification, which is useful for feature extraction when one is conducting a classification of multisource data. Lastly, it can detect outliers.

Random Forests (RF) is a modified version of previous work on CART® models (Breiman et al. 1984) and was developed by Breiman (2001a). It differs from CART® in that it produces a series of classification trees and has been found to produce improved classification accuracy (Gislason et al. 2006). It also adds another layer of randomness to the bagging process (Liaw and Wiener 2002), which turns out to be an improvement over other highly accurate classifiers, such as discriminant analysis, support vector machines (SVM), and neural networks and is not subject to overfitting (Breiman 2001a *in* Liaw and Wiener 2002).

RF can accommodate large datasets, including thousands of input variables, and it determines which variables are important in classification and gives them measures of importance (e.g., the ‘mean decrease Gini index’); RF

also provides an assessment of the classification accuracy using an out-of-bag accuracy assessment (Breiman and Cutler, n.d.). Cutler et al. (2007) provides an explanation of the Gini index and how it is used to indicate variable importance in their Appendix A. Cutler et al. (2007, p. 2783) list the “Advantages of RF compared to other statistical classifiers include (1) very high classification accuracy; (2) a novel method of determining variable importance; (3) ability to model complex interactions among predictor variables; (4) flexibility to perform several types of statistical data analysis, including regression, classification, survival analysis, and unsupervised learning; and (5) an algorithm for imputing missing values.” Furthermore, no assumptions are made about the distribution of the input data (e.g., data does not need to be normally distributed) and it can be categorical or continuous.

RF has been used in in wide range of studies and fields, from medical studies to distinguishing amongst author names in academic publications (e.g., Treeratpituk and Giles 2009) to ecological studies, and the advantages and high classification accuracy results obtained using RF are stated over and over. Pal (2005) compared RF to SVM, and while comparable in accuracy, RF had several advantages over SVM: it only required 2 input parameters, it had the ability to handle categorical data as well as deal with data issues (unbalanced and/or missing values), and it can detect outliers. Their study used Landsat data to map seven classes of land cover in the UK. RF has also been evaluated for use in ecological applications, which often contain highly dimensional data with complex interactions and missing values. Prasad et al. (2006) tested four statistical classification methods for mapping vegetation under different climatic scenarios. They conclude that RF worked very well to predict the current and potential future distributions of tree species. Similarly, Cutler et al. (2007) compared RF to four commonly used classifiers and found that it is very well suited to work with complex ecological data. While there may be differing opinions in the statistical world (e.g., see Breiman 2001b), the consensus in the literatures points to RF as

an efficient and accurate classifier for many different types of applications in various disciplines.

References for this Chapter are included in the Complete Bibliography beginning on page 179.

Chapter 2. Mapping trees outside forests using high-resolution aerial imagery: a comparison of pixel- and object-based classification approaches

Dacia M. Meneguzzo, Greg C. Liknes, Mark D. Nelson

Environ. Monit. Assess. DOI 10.1007/s10661-012-3022-1

2.1 Synopsis

Discrete trees and small groups of trees in nonforest settings are considered an essential resource around the world and are collectively referred to as trees outside forests (ToF). ToF provide important functions in the landscape, such as protecting soil and water resources, providing wildlife habitat, and improving farmstead energy efficiency and aesthetics. Despite their significance, forest and other natural resource inventory programs and geospatial land cover datasets that are available at a national scale do not include comprehensive information regarding ToF in the United States. Additional ground-based data collection and acquisition of specialized imagery to inventory these resources are expensive alternatives. As a potential solution, we identified two remote sensing-based approaches that use free high-resolution aerial imagery from the National Agriculture Imagery Program (NAIP) to map all tree cover in an agriculturally-dominant landscape. We compared the results obtained using an unsupervised per-pixel classifier (independent component analysis – [ICA]) and an object-based image analysis (OBIA) procedure in Steele County, Minnesota, USA. Three types of accuracy assessments were used to evaluate how each method performed in terms of: 1) producing a county-level estimate of total tree-covered area, 2) correctly locating tree cover on the ground, and 3) how tree cover patch metrics computed from the output compared to those delineated by a human photo interpreter. Both approaches were found to be viable for mapping tree

cover over a broad spatial extent and could serve to supplement ground-based inventory data. The ICA approach produced an estimate of total tree cover more similar to the photo-interpreted result, but the output from the OBIA method was more realistic in terms of describing the actual observed spatial pattern of tree cover.

2.2 Introduction

Trees outside forests (ToF) are considered an important land use feature in a global context and have now been included as an attribute of interest in the United Nations' Global Forest Resource Assessment. By definition, ToF are "trees on land not defined as forest and other wooded land" (FAO 2001); examples include trees that occur on agricultural and grazed lands, along waterbodies and roads, and in residential and urban settings (Rawat et al., 2003). In large portions of the central United States where agriculture dominates the landscape, tree cover exists primarily as ToF. Although scarce in terms of overall coverage, ToF provide a variety of ecological benefits, including protecting soil and water resources, providing wildlife habitat, and improving farmstead energy efficiency and aesthetics (Rietveld and Irwin 1996) and providing biomass for carbon sequestration (Schoeneberger 2005; Kort and Turnock 1999).

While the importance of ToF is recognized, a continual inventory and monitoring program for the resource does not exist in the United States. National Forest Inventories (NFIs) typically rely on minimum size and density requirements to define forests and thus do not collect information on ToF. For example, the Forest Inventory and Analysis (FIA) program of the U.S. Department of Agriculture Forest Service defines forest to be land with a minimum of 10% tree cover (or equivalent stocking) and is at least 1 acre in size (USDA Forest Service 2010). Furthermore, the area must be at least 120 feet, or 36.6 meters, in width, thus excluding narrow tree plantings, trees in urban settings, and many naturally-occurring tree corridors along streams. A study by

Perry et al. (2009) found that the estimate of total tree-covered area would exceed the estimate of forestland by at least 25% in the Great Plains region if tree resources such as ToF were included in the FIA inventory.

We do note, however, that information has been collected periodically on ToF in the United States for limited geographic areas (Hartong and Moessner (1956) in Iowa; Hansen (1985) in Kansas; Lister et al. (2009) in the Great Plains). While each of these studies relied on aerial photography and/or ground-based sampling specifically targeted at ToF, several efforts have been made to use satellite imagery to comprehensively map land cover across the conterminous U.S. (e.g., the National Land Cover Dataset (NLCD 2006) (Xian et al. 2009)). In these cases, the sensors used are too coarse to discern small groups or narrow tree plantings and do not provide consistent estimates of total tree cover (Perry et al. 2009; Liknes et al. 2010). In contrast, digital aerial imagery is typically collected at a very high spatial resolution (e.g., ≤1 meter) and is sufficient to capture small patches of trees and even individual tree crowns. The resolution, however, presents a challenge since a higher spatial resolution leads to increased spectral variation of landscape features, which makes it more difficult to statistically separate classes using traditional pixel-based classification methods and thus reduces classification accuracy; this is known as the 'H-resolution problem' (Woodcock and Strahler 1987; Marceau et al. 1990 *in* Hay et al. 1996). As such, the challenge warrants the development of new methodologies for working with this type of imagery. Two more recent options are object-based image analysis (OBIA) and independent component analysis (ICA).

Image segmentation and classification are the two main components of OBIA approaches. Segmentation is the process used to divide the imagery into homogeneous image segments, or objects, which become the processing units that are subsequently classified rather than the individual pixels (Benz et al. 2004). The image segments are groups of similar, adjacent pixels formed to represent the landscape features of interest (e.g., agricultural fields, houses, roads). User-defined settings of shape, color, compactness, and scale parameter

determine what that resulting image objects will look like. The scale parameter is a unitless number that sets the degree of heterogeneity within the image objects, so a larger scale parameter will result in larger, more heterogeneous image objects (Laliberte et al. 2007; Benz et al. 2004). The method is different from classic pixel-based procedures that rely solely on the pixel spectral values represented by digital numbers (DNs) to classify each pixel individually. OBIA procedures offer several fundamental advantages over per-pixel approaches: 1) image objects can be created at various scales (e.g., from a single tree crown to groups of trees) (de Jong and van der Meer 2004; Hay et al. 2003), 2) the use of image objects alleviates the salt-and-pepper effect often encountered in pixel-based classifications (Yu et al. 2006), and 3) numerous attributes can be obtained from image objects, including statistics such as mean and standard deviation using the DNs (Chubey et al. 2006). In addition, classification results based on image objects have been found to be more accurate than those from pixel-based procedures (Benz et al. 2004; Yu et al. 2006; Platt and Rapoza 2008; Myint et al. 2011).

While the OBIA approach has been found to produce more accurate classification results, standard OBIA-specific accuracy assessment procedures are lacking (Drăguţ and Blaschke 2006) and this remains a “hot” research topic within OBIA (Blaschke 2010). Persello and Bruzzone (2010) suggest an accuracy assessment approach “that is based on the analysis of two families of indices: 1) the traditional thematic accuracy indices and 2) a set of novel geometric indices that model different geometric properties of the objects recognized in the map.” However, this does not appear to be widely implemented at this time. The common practice found in the literature is the continued use of accuracy assessment methods that were developed for per-pixel methods, including error, or confusion, matrices and the use of descriptive statistics, such as user’s and producers accuracies (e.g., Congalton 1991). In addition, many OBIA studies use a stratified random (or proportional) sampling to select points, plots, or objects from which to create the reference data set (e.g., Myeong et al. 2003; Laliberte et

al. 2007; Johansen et al. 2007; Zhou et al. 2009; Myint et al. 2011; Peña-Barragán et al. 2011).

Although OBIA techniques such as 'Extraction and Classification of Homogeneous Objects' (ECHO) have been in existence for more than 30 years (Kettig and Landgrebe 1976), their use in extracting information from high-resolution imagery has increased markedly during the last decade; this is coincident with the increase in availability of such imagery from both satellite and aerial platforms (see Blaschke (2010) for a thorough discussion of historical and more recent OBIA research). There are numerous studies where OBIA procedures were used to produce output classifications related to natural resources. A review of the literature reveals that high resolution imagery and OBIA have been used in conjunction as an approach for mapping woody plant features in agricultural and other rural landscapes around the globe. For example, Tansey et al. (2009) used OBIA to accurately identify hedgerows with a 2-meter minimum width from aerial imagery in Berkshire, UK. Aksoy et al. (2010) carried out a study in Germany, the Czech Republic, and Cyprus in which an object-based methodology was used to automate the process of identifying linear wooded strips in agricultural areas. Other related studies include juniper cover estimation from NAIP imagery in Idaho, USA (Davies et al. 2010) and Wiseman et al. (2009) where large shelterbelts were mapped over an area covering approximately 25,900 ha using very high-resolution (62.5 cm) aerial imagery in Manitoba, Canada. The authors concluded that an object-based method was very efficient for broad-scale inventorying of shelterbelts. However, these studies were confined to small areas while more recent examples that use OBIA techniques occur over much larger areas, such as 289,755 ha in North Dakota, USA (Liknes et al. 2010) and 177,000 ha in California, USA (Peña-Barragán et al. 2011). The recent use of OBIA techniques offers a promising solution to the challenge of mapping fine-scale tree features from digital aerial imagery over a large spatial extent in the central United States.

Unlike OBIA, ICA, which is a pixel-based classification approach, is a less conventional technique that reduces the dimensionality of the input data. It was developed as a type of blind source separation (Common 1994; Hyvärinen and Oja 2000) whereby input signals could be separated into source signals without any knowledge of the original inputs. Recently, ICA has been used for unsupervised classification (Shah 2007b) and pan sharpening (Chen et al. 2011), and it has been implemented in ERDAS IMAGINE® (Shah 2007a), a popular image processing software package.

ICA is often compared and contrasted with the more well-known Principal Component Analysis (PCA). One major difference is the order of statistics used; that is, ICA makes no assumption that original source components follow a Gaussian distribution and uses skewness and kurtosis to determine the independence of input sources. The two data reduction methods are often compared; for example, Wang and Chang (2006) found that ICA-based dimensionality reduction outperformed PCA-based methodology when used with AVIRIS and HYDICE hyperspectral image data, and ICA has been used for land cover classification in the State of Iowa (e.g., ftp://ftp.igsb.uiowa.edu/gis_library/counties/lyon/HRLC_2007_60/HRLC_2007_60.html#7).

Given the need for more comprehensive information regarding ToF, the objective of the study was to investigate the aforementioned approaches as potential solutions for broad-scale mapping of all tree cover (ToF and forest) in agricultural landscapes from high resolution aerial imagery. The results offer a means for supplementing NFIs by providing information on the extent of ToF with a particular focus on methods that are efficient and at least partially automatable so that the mapping process could become a recurring part of an NFI and therefore serve to monitor trends in ToF.

2.3 Materials and Methods

2.3.1 Study Area

Steele County, located in southern Minnesota, USA, was selected as the study area (Figure 2.1). The county is nearly 111,000 ha in size and the landscape is similar to that found throughout the central United States. The dominant landscape feature is row-crop agriculture, and other cover types include trees, farmsteads, urban development and roads, rivers and lakes. The city of Owatonna is the county seat and about two-thirds of the county's population resides there. The non-urban portion of the county is comprised of 934 farms according to the 2007 Census of Agriculture (USDA National Agricultural Statistics Service 2009).

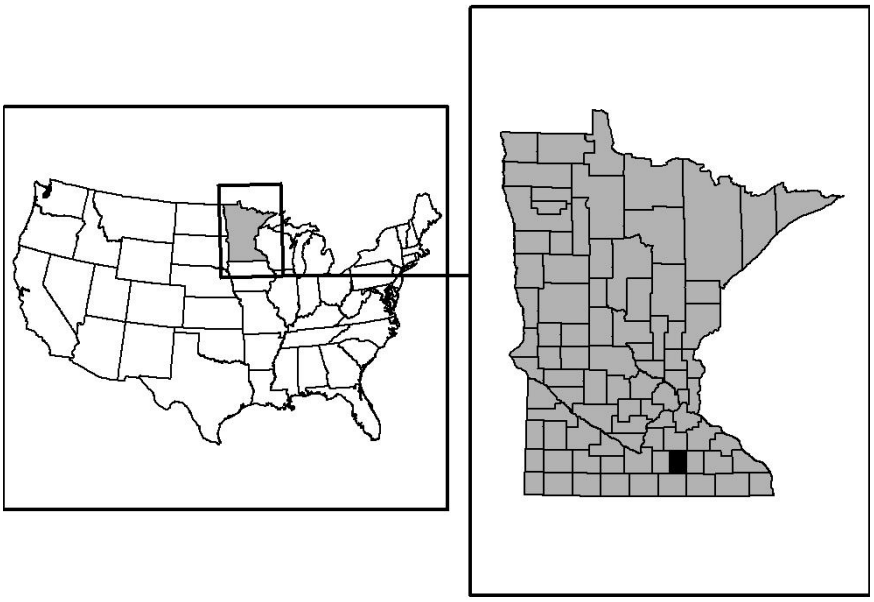


Figure 2.1. Location of Steele County in southern Minnesota, USA.

2.3.2 High Resolution Imagery

Digital aerial imagery from the U.S. Department of Agriculture's Farm Service Agency National Agriculture Imagery Program (NAIP) was obtained for this study. NAIP imagery is collected during the growing season (leaf-on) primarily for agricultural compliance monitoring and has been captured on a

routine basis since 2003. The return interval varies by state, and datasets from 2003 (1-m), 2004 (2-m), 2005 (2-m), 2006 (2-m), 2008 (1-m), 2009 (1-m), and 2010 (1-m) exist for the state of Minnesota. Data from 2008 were used because of the availability of the near-infrared (NIR) band in addition to the normally-acquired red, green, and blue bands. Images were obtained in uncompressed TIFF format and had been divided into a series of forty-nine tiles with 300-m of overlap between adjacent images. The tiles have a variety of image acquisition dates from throughout the growing season (June, July, or August).

Input data layers for the OBIA and ICA approaches were obtained or derived from the NAIP imagery and included the red, green, blue, and NIR spectral bands, the normalized difference vegetation index (NDVI), and a green texture band. NDVI is derived using the NIR and red bands from the imagery where the difference between the two is divided by the sum of the two bands. The index is commonly used for identifying vegetation and can be used for other purposes such as identifying stressed versus healthy vegetation (Tucker and Choudhury 1987). However, its use in this study was simply to add other useful information for detecting tree cover. While NDVI is useful for identifying vegetation in general, trees needed to be discriminated from other surrounding vegetation, so the use of texture layers was incorporated. Texture is a way to measure the visual roughness versus smoothness of features in an image (Haralick et al. 1973; Lillesand and Kiefer 1994) and is important for distinguishing tree cover from other vegetation, such as grassy lawns (Zhang 2001 *in* Tansey et al. 2009; Myeong et al. 2003).

2.3.3 OBIA Approach

The workflow for the OBIA approach is shown in Figure 2.2. The segmentation and classification routines were carried out using eCognition Developer software v. 8.0.1 (Definiens 2010). A trial-and-error approach and visual inspection of the results was employed in order to determine which user-defined settings produced the most meaningful image objects in order to meet

our study objectives. To begin, the ‘multiresolution segmentation’ algorithm was used to segment each image tile into fine-scale image objects (called “Level 1”) with the following settings: scale parameter = 15, shape = 0.2, and compactness = 0.9 (Figure 2.3). A second level (“Level 2”) of larger image objects was then created using a scale parameter of 20 and the Level 1 image objects as building blocks. All subsequent processing occurred on the Level 2 image objects. This was done as means to reduce the large number of image objects for more efficient successive processing. More emphasis was given to color (0.8) rather than shape (0.2) during the segmentation processes and only the 4 spectral bands were utilized to create the Level 1 and Level 2 image objects. Compactness was set high (0.9) so more circular-shaped image objects were created in an attempt to accurately represent the shape of tree crowns. The NDVI and texture information were used in subsequent steps for separating tree from no-tree image objects. The segmentation/classification routine was developed for one image tile and then applied to the remaining 48 images using a programming script to automate the processing.

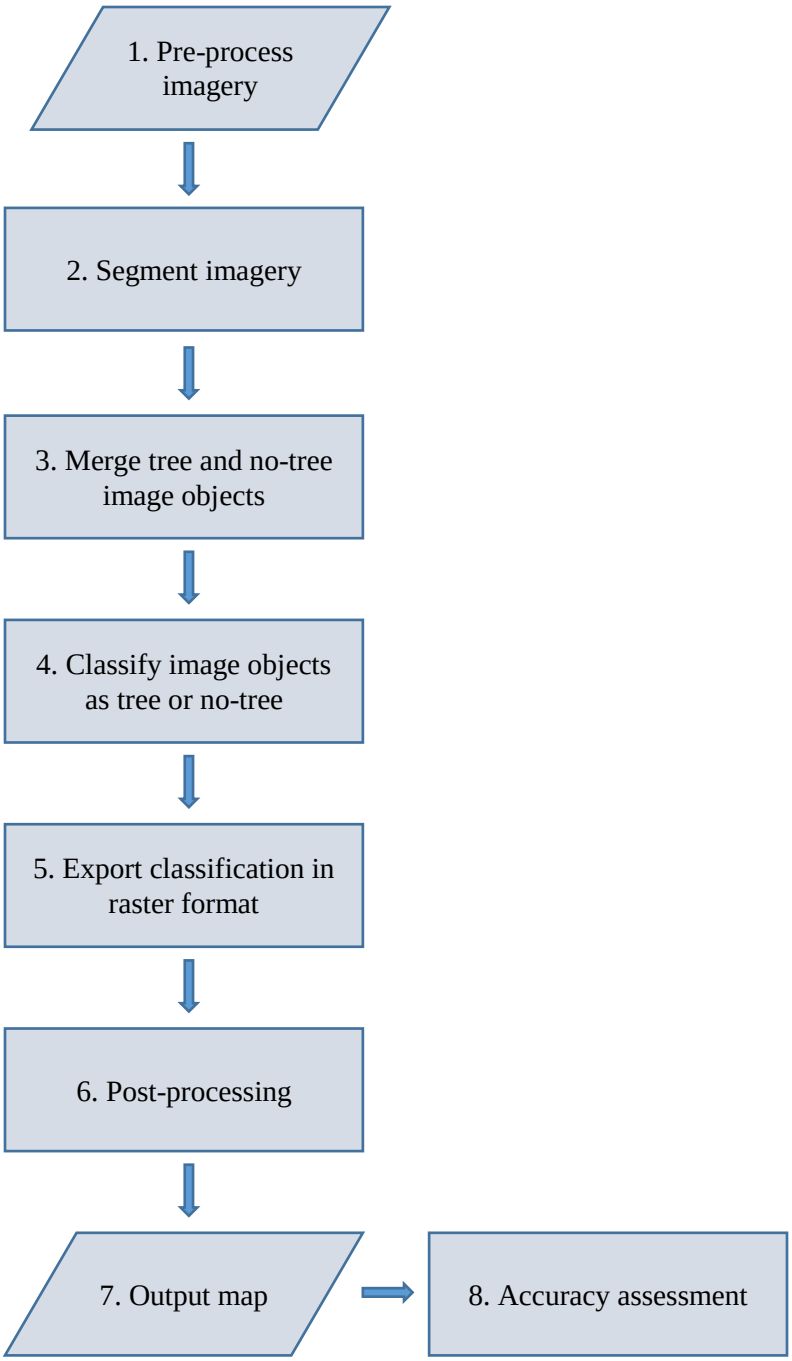


Figure 2.2. Object-based image analysis workflow for classifying NAIP imagery into tree and no-tree classes. ArcGIS® and ERDAS IMAGINE® software were used in Step 1 to pre-process the imagery. Steps 2 through 5 run sequentially in eCognition software and processing multiple images was automated using a programming script. Steps 7and 8 were carried out using a python™ script.

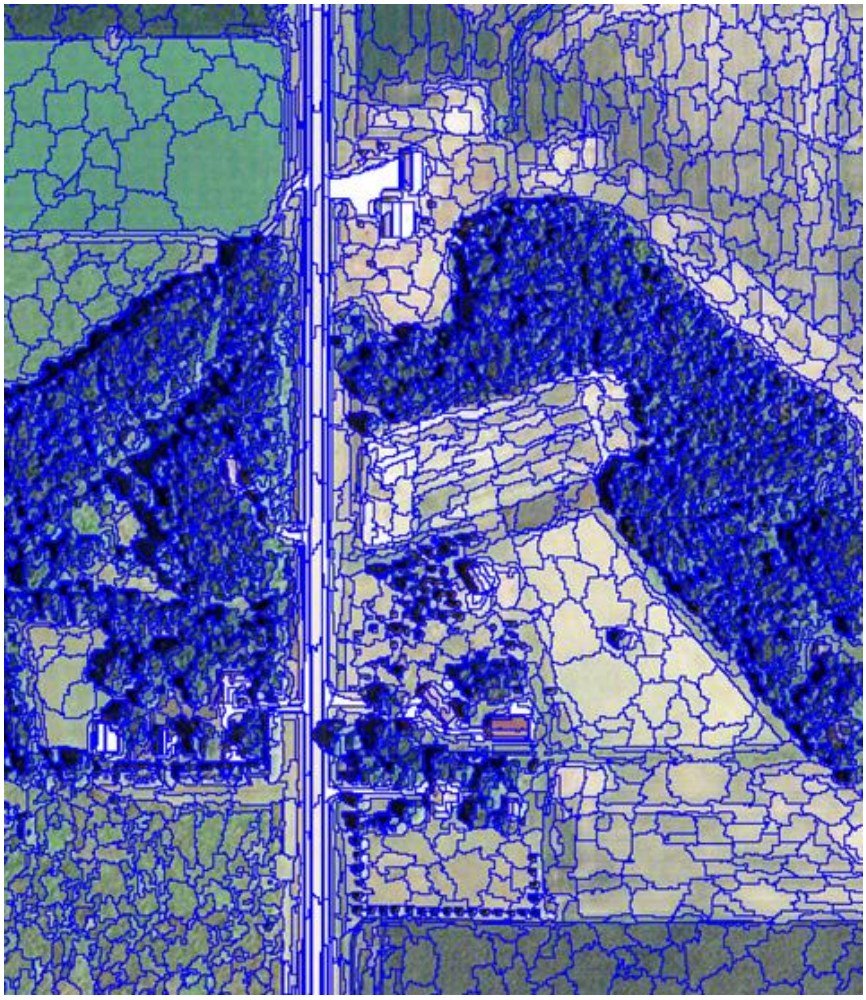


Figure 2.3. Example of Level 1 image objects created in eCognition from NAIP imagery in Steele County, MN.

The primary goal during segmentation was to maintain image objects that were purely tree canopy, whether it was a single tree crown or continuous canopy. This was accomplished using a series of thresholds and an increasing scale parameter to iteratively merge the no-tree objects into larger and larger objects (e.g., farm fields) by capitalizing on the NDVI and texture information. The tree image objects were also aggregated using the 'multiresolution segmentation region grow' algorithm to make larger, more continuous canopy objects and reduce the total number of image objects. Typically, there would be

more than 200,000 Level 1 image objects reduced to about 2,000 per image tile using this process (Figure 2.4a). There are many image object attributes (spectral, spatial, and textural) that can be incorporated during the segmentation and classification processes. In this study, the processes relied primarily on the following attributes (using the mean value of the image object) to distinguish tree cover from the no-tree image objects: brightness (combined value of the red, green, and NIR input bands), texture of the green band and NDVI, and values of each of the four spectral bands.

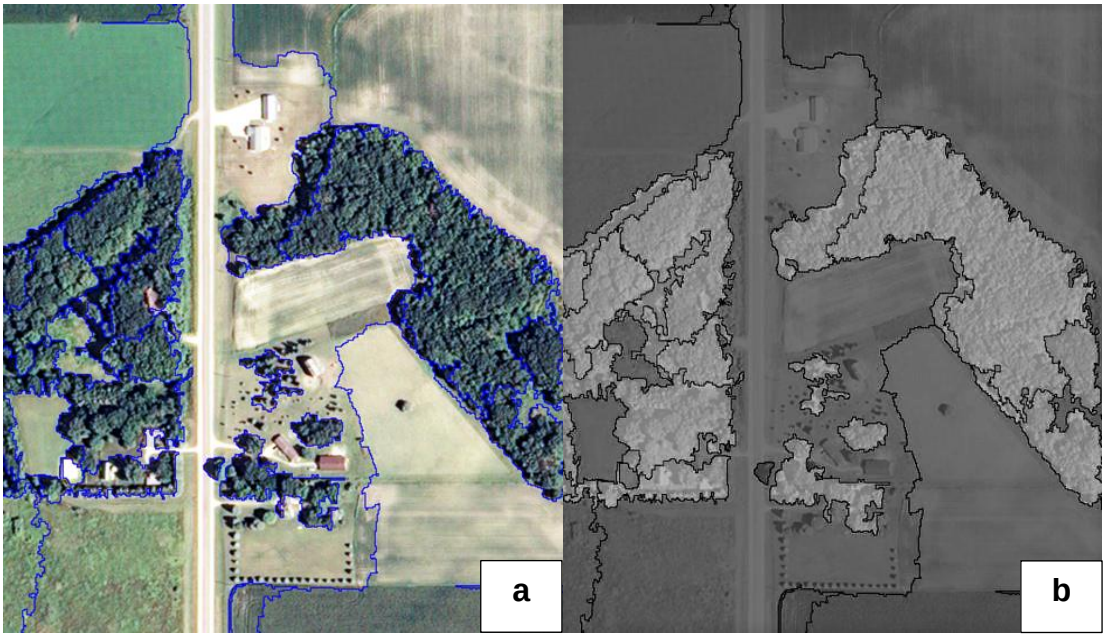


Figure 2.4. Final level of image objects (a) and classification of final level of image objects created in eCognition from NAIP imagery in Steele County, MN; 'tree' objects are represented in light gray (b).

During the classification phase, image objects were assigned to one of two classes: tree or no-tree (Figure 2.4b). Classification rules were developed using mostly the features listed in the previous paragraph. Threshold were developed using observed feature information for no-tree objects compared to that of tree objects. We used lower mean values of red, green, blue, and brightness, and higher mean NDVI, NIR and green texture values to distinguish

tree cover from the surrounding areas. Classification results for each image tile were exported in raster format. ArcGIS® Desktop v.9.3.1 software (ESRI Inc. 2009) was used to mosaic the raster outputs, clip the compiled output to the county boundary and convert it to vector format. Lastly, the estimate of total tree-covered area for the county was obtained by calculating the area of the 'tree' class from the final output.

2.3.4 ICA Approach

For the second mapping approach, a workflow (Figure 2.5) was implemented in which an NDVI image band was created from the NAIP imagery as well as a green texture band in step 1. Median filters have been used for noise reduction with high resolution imagery (e.g., Mora et al. 2010), so a 5x5 median filter was also applied in step 1 to all input bands; these filtered bands were used as inputs into the ICA data reduction step (step 2). At this point, the ICA output bands were inspected (step 2b) to determine if useful information was contained. Although it was rare, there were cases where ICA components contained no information and appeared as a blank image to the interpreter. These particular bands were removed from further processing. Next, the ICA bands were clustered using ISODATA in step 3 into 20 classes and a Maximum Likelihood Classifier was used to assign the clusters into a class. An interpreter then examined the 20 classes in conjunction with the original NAIP imagery and selected those classes that best represented tree cover (step 4b). Once the class labels (i.e., tree and no-tree) were assigned, a minimum mapping unit (MMU) of 20 pixels, or 20 m^2, was applied in step 5. This particular MMU was chosen because it represented a conservative minimum size for a single tree crown. Similar to the OBIA approach, the post-processing procedures in step 5 included mosaicking the classified output tiles together, clipping the compiled output to the county boundary, and converting the final raster to vector format using the same software. Again, the total area of tree cover was obtained from the county-level classified output. While steps 2b and 4b required human intervention, all other

steps are fully automated, and batch processing was used for all 49 image tiles in the study area.

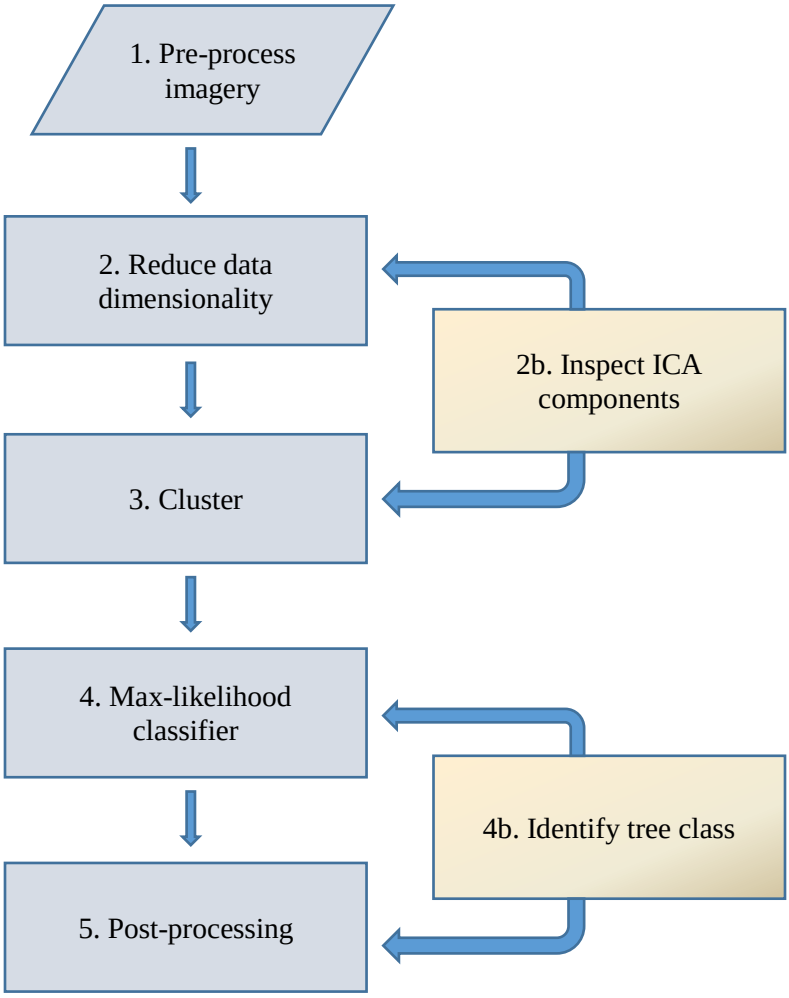


Figure 2.5. Unsupervised classification workflow for classifying NAIP imagery into tree and no-tree classes. Steps 2b and 4b are manual while all other steps are fully automated and can be batch processed for many images

2.3.5 Accuracy Assessment

Three different accuracy assessments were conducted in this study: nonsite-specific, site-specific, and a targeted assessment. Because the goal of

the study was to obtain county-level area estimates of tree cover and to compare the estimates to sample-based FIA forest area estimates, we began with a nonsite-specific, area-based accuracy assessment. Area-based assessments are typically used to determine map accuracy by first aggregating units (pixels or image objects) to a larger area (Lunetta and Lyon 2004). While area-based assessments have inherent drawbacks (e.g., Congalton 1991), the approach met the evaluation objective of the study. The area-based assessment consisted of a cluster sample framework and heads-up digitizing to estimate the total area of tree cover across the county. Cluster sampling allowed us to sample/delineate tree cover within smaller units, or blocks, placed throughout the county rather than digitizing all tree cover, which would be very time-intensive and expensive.

In order to supplement the area-based assessment with spatially-explicit information about map accuracy, we conducted site-specific and targeted assessments of the OBIA and ICA methods as well. The site-specific accuracy assessment was employed to evaluate the locational accuracy of the two thematic classes (tree and no-tree) compared to the reference data. This type of assessment is important because it considers the location of each class, not only the total area (Jensen 1996). Within the targeted assessment, a variety of landscape pattern metrics were compared across three different landscape types within the Steele County study area: agricultural, riparian, and urban. Because the spatial arrangement of tree cover in the landscapes relates to ecosystem processes, it was appropriate to examine the consistency of metrics derived from the ICA and OBIA outputs that were used to quantify tree-cover patterns.

2.3.5.1 Area-based Accuracy Assessment

The area-based accuracy assessment was designed to determine how well the OBIA and ICA methods performed with regard to correctly estimating the proportion of tree cover in the county. Specifically, a cluster sample was employed for the study area using equal-sized grids (1 km^2) with centers separated by 3 km (Figure 2.6), and resulted in a total of 108 grids. For each grid

square, a trained photo interpreter used heads-up digitizing methodology to map tree cover. An estimate of the proportion of tree cover in the county was then calculated by

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} p_i$$

where n is the total number of grid squares in the sample and p_i is the proportion of tree cover for the i^{th} grid square (adapted from Thompson 2002).

The standard error is given by

$$SE(\hat{p}) = \sqrt{\frac{s^2}{n}(1 - \frac{n}{N})}$$

Where N is the number of km^2 in the study area and s^2, the sample standard deviation is given by

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(p_i - \hat{p})^2$$

The county-level proportion estimate of tree cover is easily converted to an areal unit by multiplying the estimate by the total area of the county.

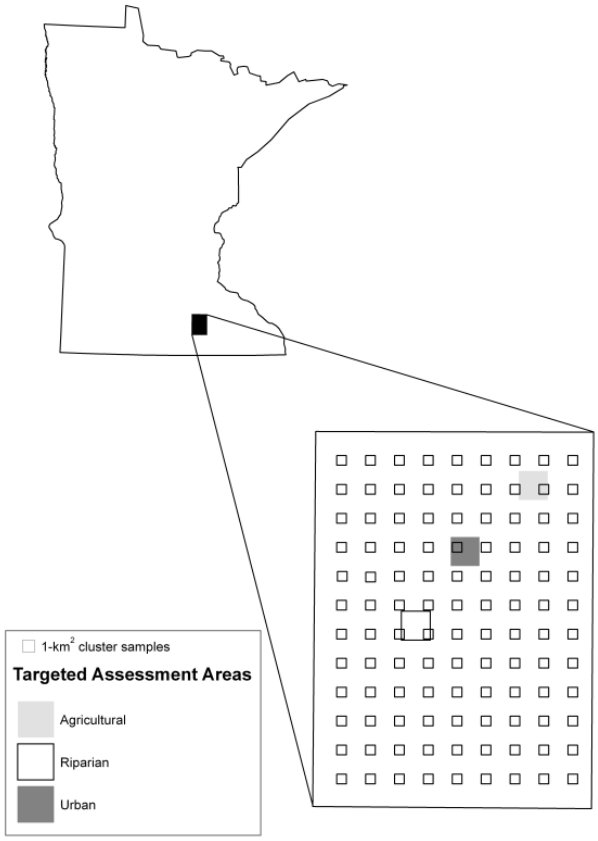


Figure 2.6. Cluster sample design of 1 km x 1 km blocks at 3 km intervals for the area-based assessment and targeted landscape assessment types (agricultural, riparian, and urban).

2.3.5.2 Site-specific Accuracy Assessment

Site-specific accuracy assessments were used to directly compare the classified output derived from the ICA and OBIA approaches to the reference data, and the accuracy of each approach was represented in an error matrix. This type of accuracy assessment is more complete than nonsite-specific assessments because it accounts for the locational accuracy of the classified output and not only the total area.

An important part of the site-specific assessment is collecting unbiased reference data to which the classified output is compared. To accomplish this, we used stratified random sampling to select 50 samples from each stratum (i.e., the tree and no-tree classes) to ensure that both were adequately represented in the accuracy assessment (Congalton 1991). Using the high-resolution NAIP imagery as the reference data source, a trained photo interpreter labeled each sample as 'tree' or 'no-tree' and these were compared to the output classification derived from the ICA and OBIA approaches. The agreement/disagreement results were summarized in error matrices and descriptive statistics including producer's accuracy (measure of omission error), user's accuracy (measure of commission error), and overall accuracy were calculated to describe the accuracy of each thematic class produced by both classification methods.

2.3.5.3 Targeted Assessment

Tree cover in the three different types of landscapes was delineated into patches using heads-up digitizing to facilitate a more detailed comparison of the classification results between the OBIA and ICA methods. Grid squares (3km x 3km) were placed in a riparian area, an agricultural area with windbreaks, and in an urban setting (Figure 6). For each of these areas, selected patch metrics were calculated using the Patch Analyst Extension in ArcMap™ and are listed in Table 2. The targeted assessment further characterized how each approach performed in characterizing tree cover in different landscape types and addresses the questions: 1) is each approach equally applicable in all landscapes? And, 2) how do the approaches perform in terms of producing spatially accurate information in the various landscapes?

Area, number of patches, average patch size, median patch size, and patch density metrics are standard measures in landscape-level analyses that describe the amount and spatial arrangement of patches, which are patches of tree cover in this study. For example, a large number of patches, a small average patch size, and a high patch density indicate that tree cover in the landscape is

fragmented, occurring as many small, separate patches. Mean perimeter-area ratio is a measure used to describe the average patch shape in the landscape. The mean is calculated by summing the perimeter-area ratio of each patch and dividing by the total number of patches. A higher mean perimeter-area ratio indicates that, on average, the patches are more complex and irregular in shape. While perimeter-area ratio is a common and simple way to indicate shape, it is influenced by the size of the patch. Mean patch fractal dimension, however, allows patches to be weighted by size to help correct this problem; a value close to 1 indicates that patches have simple boundaries regardless of size whereas a value near 2 means that the patch shapes are more complex across various patch sizes (McGarigal and Marks 1995).

2.4 Results

2.4.1 Tree-covered Area

Total tree-covered area results from the three methods are presented in Table 2.1. The area of tree cover found using the ICA approach was very similar (9% difference) to the estimate obtained from the heads-up digitizing in the cluster sample while the result from the OBIA approach was substantially higher (53% difference). In comparison, the 2010 FIA estimate of forest land for the study area is much smaller (72% difference from the cluster sample) than all other estimates of total tree cover.

Table 2.1. Estimates of tree-covered area in Steele County, MN using three methods. Tree cover was assessed by a human photo interpreter (PI) using heads-up digitizing, semi-automated image object-based image analysis (OBIA), and unsupervised classification (ICA) approaches. Forest Inventory and Analysis (FIA) estimates of forest land were used as a comparison.

	PI	OBIA	ICA	FIA estimate
Area of tree cover (hectares)	5650	9760	5180	2669
Proportion tree cover	0.051	0.088	0.047	0.017
Standard error	0.0038	*	*	0.57

*the OBIA and ICA methods are census approaches while PI and FIA are sample based.

2.4.2 Site-specific Accuracy Assessment

The results of the site-specific accuracy assessment for each approach indicate that both methods produced reliable maps of tree cover versus no-tree cover. The overall accuracy of each classification method was high, 88% and 95% for the ICA and OBIA approaches, respectively. The producer's and user's accuracies for the classification output derived from the OBIA approach were above 90% for both classes. The accuracy assessment results for the ICA approach were more varied. While the user's accuracy for the 'tree' class was 100%, the producer's accuracy was only 76%. For the 'no-tree' class, user's accuracy was 81% and the producer's accuracy was 100%.

2.4.3 Tree Cover Patch Metrics

The three approaches performed differently in providing estimates of tree cover patch metrics (Table 2.2). The ICA approach resulted in a smaller tree-covered area relative to the heads-up digitized approach while the OBIA method resulted in more tree-covered area for all three landscapes. The ICA method also

tended to produce many smaller patches of tree cover in all landscapes compared to the other two approaches. This is illustrated by the substantially higher numbers of patches, smaller average and median patch sizes, and higher patch densities estimated using the ICA method. Regarding patch shape, the ICA approach produced patches with more complex perimeters, as indicated by mean patch fractal dimension values close to 2, in all the target landscape types.

Table 2.2. Comparison of landscape metrics from a targeted assessment of tree cover in agricultural, riparian, and urban landscapes (each 3km x 3km) in Steele County, MN using three methods. Tree cover was assessed by a human photo interpreter (PI) using heads-up digitizing, semi-automated image object-based image analysis (OBIA), and unsupervised classification (ICA) approaches.

Metric	Agricultural			Riparian			Urban		
	PI	OBIA	ICA	PI	OBIA	ICA	PI	OBIA	ICA
Tree-covered area (ha)	20.5	27.4	14.6	100.7	129.1	77.8	237.5	301.2	176.8
Number of patches	93	49	919	105	81	1880	3967	169	12535
Average patch size (ha)	0.22	0.56	0.02	0.96	1.59	0.04	0.06	1.78	0.01
Median patch size (ha)	0.05	0.25	0.005	0.03	0.29	0.004	0.01	0.45	0.004
Standard deviation of patch size (ha)	0.43	0.67	0.07	5.82	8.22	0.90	0.51	3.98	0.09
Patch density (patches/km^2)	10	5	102	12	9	209	441	19	1393
Mean perimeter-area ratio	2695	2149	10622	3116	2576	11740	4611	2526	12628
Mean patch fractal dimension	1.53	1.57	1.96	1.55	1.61	1.98	1.63	1.56	1.99

In the agricultural and riparian landscapes, the metric results from the OBIA approach were consistently more similar to the PI results than those derived from the ICA output. Mean perimeter-area ratio and mean patch fractal dimension metrics from the PI and OBIA methods are comparable, indicating that the two approaches tended to produce similarly-shaped patches of tree cover. In contrast, the ICA approach created patches with much more complex shapes,

e.g., often small and blocky in shape with a high number of edges per patch. Figure 2.7(a-d) shows an example from the riparian target area (a) comparing the output from heads-up digitizing (b), and the OBIA (c) and ICA (d) approaches.

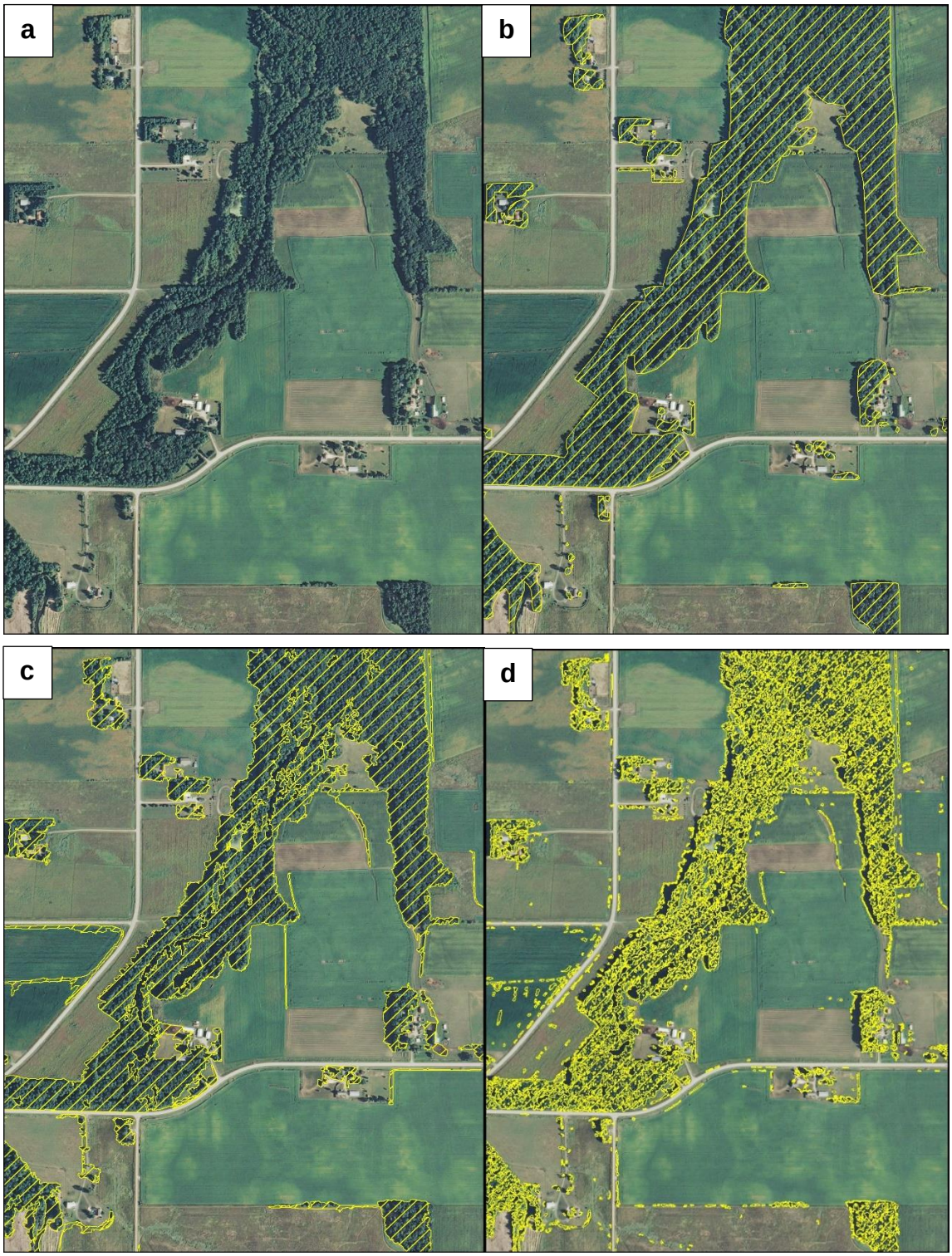


Figure 2.7. Example of the riparian target type and comparisons of delineation of tree cover using three methods in Steele County, MN. Unclassified NAIP image (a), tree cover assessed by a human photo interpreter (PI) using heads-up digitizing (b), semi-automated image object-based image analysis (OBIA) (c) and unsupervised classification (ICA) (d) approaches.

The metric results for the urban landscape varied widely among the three methods. In contrast to the ICA approach, the OBIA method produced far fewer and larger patches: 169 compared to more than 12,000, with an average size of 1.78 ha versus 0.01 ha. The only OBIA metrics that were somewhat similar to the PI results were tree-covered area and mean patch fractal dimension. However, the PI and ICA methods produced many smaller patches that were more similar in average size than the OBIA method. Again, the ICA approach produced the most complex-shaped patches of tree cover.

2.5 Discussion

Tree cover in nonforest settings is a sparse yet important resource. The lack of current inventory and monitoring programs of ToF is a concern; however, obtaining accurate information about its extent and location is challenging. Additional ground-based data collection as part of an NFI is relatively expensive and does not provide detailed spatial information that can be used in other research and applications, such as determining ecosystem function. Commonly used and widely available land cover datasets are acquired at spatial resolutions that are too coarse to detect small patches and narrow bands of tree cover and higher resolution imagery can be costly. In order to find a potential solution to these issues, two remote-sensing based approaches, OBIA and ICA, were examined for mapping trees in an agricultural landscape using freely available, very high resolution (1-m) aerial imagery.

When determining the total area of tree cover for Steele County, Minnesota, the ICA approach produced an estimate similar to that found from

heads-up digitizing in the cluster sample (5,180 ha versus 5,650 ha respectively). The OBIA estimate, on the other hand, was much higher (9,760 ha), and all tree cover estimates are substantially higher than the FIA estimate of forest land (2,669 ha). The findings reinforce what other authors have reported, that the definition of forest land excludes a significant portion of tree cover in agricultural landscapes from NFIs. In this case, total tree cover in Steele County is potentially 3 times as much as the forest land area estimate would indicate if we consider the PI estimate to be the most accurate reference or standard. Examination of the OBIA and ICA outputs indicate that both methods struggle with shadows and grassy vegetation along roadways and in ditches, and sometimes erroneously label these areas as tree cover. The speckled appearance of the ICA output is due to the occurrence of many tiny, disjunct patches and is evidenced by the results shown in Table 2.2 (e.g., large numbers of patches with very small average patch sizes) and Figure 7d. The OBIA approach alleviated this problem by aggregating pixels into larger image objects and prevented the formation of such extraneous tiny patches. However, this contributed to the overestimation of tree-covered area when image objects were misclassified. For example, inspection of the OBIA output revealed that some farm fields and wetland areas were misclassified and resulted in additional large patches of tree cover. Additional research and development on the procedures will help correct the shortcomings.

While the accuracy of the OBIA and ICA methods can be improved, they do provide an advantage over the cluster sample with regard to providing more spatial detail. The arrangement of tree cover and its proximity to other land cover features provide information about the ecosystem function. For example, tree cover arranged in winding, narrow strips in proximity to streams serve as riparian buffers. Tree cover arranged in rectangular or L-shaped blocks in proximity to buildings offers protection from the weather and increases energy efficiency in those structures while linear strips of trees along field edges provide shelter from the wind and help prevent soil erosion. If a spatial database of tree cover was

constructed using the OBIA or ICA approaches, ecosystem function information could be extracted, and that is not an option readily achievable using data from a PI cluster sample.

The targeted assessment provides additional information about the ability of each method to characterize the spatial arrangement of tree cover compared to patches delineated by a human photo interpreter. In this case, it is easy to see that the pixel-based approach (ICA) leads to an extremely high estimate of the number of patches and correspondingly low average patch size. The OBIA approach is much better at mimicking how a human interpreter groups trees into patches, thus producing a result that most closely resembles that from the PI approach. While many authors focus on the processing efficiency of OBIA methods and the general better appearance of output maps and export options (e.g., Benz et al. 2004), this study points to another potential advantage over pixel-based approaches: the ability to produce better patch-based metrics for describing spatial pattern. However, when conducting ecological studies that use landscape metrics, it is important to remember that the spatial resolution of the imagery used in the metric calculations will affect the results. The principal investigator of the study should carefully select metrics and imagery that are appropriate for their research objectives.

Examination of the results in the urban target area clearly indicates that this type of landscape is extremely complex and it is difficult to accurately delineate tree cover using either of the remote-sensing based approaches. The OBIA method had a tendency to group together individual tree crowns that were in close proximity to each other while a human interpreter was able to delineate each crown separately. In contrast, the ICA approach produced output that was very speckled in appearance and often misclassified shadows around buildings and tree crown edges as tree cover. The highly variable results of the metrics from the three approaches led us to conclude that a new, separate classification model needs to be developed for urban landscapes and that future work will focus strictly on rural settings. Furthermore, the similarity of the results between

the PI and OBIA approaches in the agricultural and riparian target areas is reassuring since these are the types of areas in which we are ultimately interested in for natural resource inventory and monitoring purposes. However, it is very likely that the OBIA approach will need to be modified in terms of adjusting the user-defined settings in the eCognition software when using other imagery and/or when moving to a much different geographic area in order to create the most meaningful image objects.

The need was highlighted for methodologies that can be accurately and efficiently applied to mapping tree cover in areas where the resource is not inventoried with satisfactory results. The workflows developed for both the OBIA and ICA approaches are at least partially automatable, using either programming scripts or built-in batch processing capability of the software used, and do not require the use of expensive imagery. As such, either method represents a viable approach to mapping tree cover over a broad spatial extent and could serve to supplement NFIs. The utility was demonstrated on a county of more than 100,000 ha of land area and the results using two different accuracy assessment approaches were compared. Remote sensing-based approaches, such as OBIA or ICA, represent a step forward from traditional sample-based PI methods because of the additional spatial detail they provide. In addition, NAIP imagery is available on a periodic schedule so it would be possible to repeatedly monitor these tree resources over time. Because the OBIA approach produced classification results that were more accurate in terms of spatial location and also provides more reasonable information about the spatial pattern of tree cover, it is the better choice.

2.6 References

Aksoy, S., Akcay, H. G., & Wassenaar, T. (2010). Automatic mapping of linear woody vegetation features in agricultural landscapes using very high resolution imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 48(1), 511-522.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, *58*(3-4), 239-258.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing, 65,* 2-16.

Blaschke, T., & Strobl, J. (2001). What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. *GIS – Zeitschrift fur Geoinformationssysteme, 14* (6), 12-17.

Chen, F., Guan, Z., Yang, X., & Cui, W. (2011). A novel remote sensing image fusion method based on independent component analysis. *International Journal of Remote Sensing*, *32*(10), 2745-2763.

Chubey, M. S., Franklin, S. E., & Wulder, M. A. (2006). Object-based analysis of Ikonos-2 imagery for extraction of forest inventory parameters. *Photogrammetric Engineering & Remote Sensing*, *72*(4), 383-394.

Comon, P. (1994). Independent component analysis, a new concept? *Signal Processing 36*(3), 287-314.

Congalton, R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37*, 35-46.

Davies, K. W., Petersen, S. L., Johnson, D. D., Davis, D. B., Madsen, M. D., Zvirzdin, D. L., & Bates, J. D. (2010). Estimating juniper cover from National Agriculture Imagery Program (NAIP) imagery and evaluating relationships between potential cover and environmental variables. *Rangeland Ecology & Management*, *63*(6), 630-637.

Definiens AG. (2010). eCognition Developer 8.0.1 Reference Book. Retrieved from http://www.definiens.com.

De Jong, S.M., & van der Meer, F.D. (Eds.) (2004). *Remote Sensing Image Analysis: Including the Spatial Domain, vol. 5*. Dordrecht: Kluwer Academic Publishers.

Dragut, L. and T. Blaschke. (2006). Automated classification of landform elements using object-based image analysis. *Geomorphology, 81,*330-344.

ESRI Inc. (2009). ArcGIS Desktop: Release 9.3.1. Redlands, CA: Environmental Systems Research Institute.

FAO. (2001). Global forest resources assessment 2000. Main report. Rome: Food and Agriculture Organization of the United Nations.

Hansen, M.H. (1985). Line intersect sampling of wooded strips. *Forest Science*, *31*(2), 282-288.

Haralick, R.M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, *3*(6), 610-621.

Hartong, A.L., & Moessner, K.E. (1956). Wooded strips in Iowa. Forest Survey Release 21. USDA Central States Forest Experiment Station, Columbus, OH.

Hay, G. (2003). A comparison of three image-object methods for the multiscale analysis of landscape structure. *ISPRS Journal of Photogrammetry and Remote Sensing*, *57*(5-6), 327-345.

Hay, G.J., Niemann, K.O., & McLean, G.F. (1996). An object-specific image-texture analysis of H-resolution forest imagery. *Remote Sensing of Environment*, *55*(122), 108-122.

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Nerual Networks*, *13*, 411-430.

Johansen, K., Coops, N.C., Gergel, S.E., Stange, Y. (2007). Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification. *Remote Sensing of Environment, 110,* 29-44.

Kettig, R.L., & Landgrebe, D.A. (1976). Classification of multispectral image data by extraction and classification of homogeneous objects. *IEEE Transactions on Geoscience Electronics*, *GE-14*(1), 19-26.

Kort, J., & Turnock, R. (1999). Carbon reservoir and biomass in Canadian prairie shelterbelts. *Agroforestry Systems*, *44*, 175-186.

Laliberte, A. S., Fredrickson, E. L., & Rango, A. (2007). Combining Decision Trees with Hierarchical Object-oriented Image Analysis for Mapping Arid Rangelands. *Photogrammetric Engineering & Remote Sensing*, *73*(2), 197-207.

Liknes, G.C., Perry, C.H., & Meneguzzo, D.M. (2010). Assessing tree cover in agricultural landscapes using high-resolution aerial imagery. *Journal of Terrestrial Observation*, *2*(1), Article 5.

Lister, A., Scott, C., & Rasmussen, S. (2009). Inventory of trees in nonforest areas in the Great Plains states. In: McWilliams, W., Moisen, G., Czaplewski, R. (Comps.), *Forest Inventory and Analysis (FIA) Symposium 2008; October 21-23, 2008; Park City, UT*. Proc. RMRS-P-56CD. Fort Collins, CO, pp. 17:1-7.

Lunetta, R.S., Lyon, J.G. (Eds.) (2004). *Remote Sensing and GIS Accuracy Assessment*. Boca Raton: CRC Press.

Marceau, D.J., Howarth, P.J., Dubois, J.M., & Gratton, D.J. (1990). Evaluation of the grey-level co-occurrence matrix method for land-cover classification using SPOT imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *28*(4), 513- 519.

McGarigal, K., & Marks, B.J. (1995). FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure. USDA Forest Service Generap Technical Report PNW-351.

Mora, B., Wulder, M.A., & White, J.C. (2010). Segment-constrained regression tree estimation of forest stand height from very high spatial resolution panchromatic imagery over a boreal environment. *Remote Sensing of Environment*, *114*(11), 2474-2484.

Myeong, S., Nowak, D.J., Hopkins, P.F., & Brock, R.H. (2003). Urban cover mapping using digital, high-spatial resolution aerial imagery. *Urban Ecosystems, 5*, 243-246.

Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment*, *115*(5), 1145-1161.

Peña-Barragán, J.M., Ngugi, M.K., Plant, R.E., Six, J. (2011). Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sensing of Environment*, *115*(6), 1301-1316.

Perry, C.H., Woodall, C.W., Liknes, G.C., & Schoeneberger, M.M. (2009). Filling the gap: improving estimates of working tree resources in agricultural landscapes. *Agroforestry Systems*, *75*, 91-101.

Persello, C. and L. Bruzzone. (2010). A novel protocol for accuracy assessment in classification of very high resolution images. *IEEE Transactions on Geoscience and Remote Sensing, 48,* 1232-1244.

Platt, R.V., & Rapoza, L. (2008). An evaluation of an object-oriented paradigm for land use/land cover classification. *The Professional Geographer, 60*(1), 87-100.

Rawat, J.K., Dasgupta, S., Kuman, R., Kumar, A., & Chauha, K.V.S. (2003). Training manual on inventory of trees outside forests (TOF). Retrieved from http://www.fao.org/3/ac840e/AC840E00.htm#TOC

Rietveld, B., & Irwin, K. (1996). Agroforestry in the United States. Retrieved from http://www.unl.edu/nac/agroforestrynotes/an01g01.pdf

Schoeneberger, M.M. (2005). Agroforestry: working trees for sequestering carbon on ag lands. In: Brooks, K.N., Foliot, P.F. (Eds.), *Moving Agroforestry into the Mainstream. Proceedings of the 9th North American Agroforestry Conference*. St. Paul, MN, 13 pp.

Shah, C.A., Anderson, I., Gao, Z., Hao, S., Leason, A. (2007a). Towards the development of next generation remote sensing technology – ERDAS IMAGINE incorporates a higher order feature extraction technique based on ICA. In: *Proceedings of the ASPRS 2007 Annual Conference*, Bethesda, MD.

Shah, C.A., Varshney, P.K., Arora, M.K. (2007b). ICA mixture model algorithm for unsupervised classification of remote sensing imagery. *International Journal of Remote Sensing*, *28*(8), 1711-1731.

Tansey, K., Chambers, I., Anstee, A., Denniss, A., Lamb, A. (2009). Object-oriented classification of very high resolution airborne imagery for the extraction of hedgerows and field margin cover in agricultural areas. *Applied Geography, 29*, 145-157.

Thompson, S. (2002). *Sampling* (2nd ed.) New York: John Wiley and Sons, Inc.

Tucker, C.J., & Choudhury, B.J. (1987). Satellite remote sensing of drought conditions. *Remote Sensing of Environment*, *23*(2), 243-251.

USDA National Agricultural Statistics Service. (2009). 2007 Census of Agriculture County Profile – Steele County, MN. Retrieved from

http://www.agcensus.usda.gov/Publications/2007/Online_Highlights/County_ Profiles/Minnesota/cp27147.pdf

USDA Forest Service. (2010). Forest inventory and analysis nation core field guide. vol.1: field data collection procedures for phase 2 plots, version 5.0 Retrieved from http://www.nrs.fs.fed.us/fia/data-collection/

Wang, J., & Chang, C. (2006). Independent Component Analysis-based dimensionality reduction with applications in hyperspectral image analysis. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(6), 1586-1600.

Wiseman, G., Kort, J., & Walker, D. (2009). Quantification of shelterbelt characteristics using high-resolution imagery. *Agriculture, Ecosystems & Environment*, 131, 111-117.

Woodcock, C.E., & Strahler, A.H. (1987). The factor of scale in remote sensing. *Remote Sensing of Environment*, *21*, 311-332.

Xian, G., Homer, C., & Fry, J. (2009). Updating the 2001 National Land Cover Database land cover classification to 2006 by using Landsat imagery change detection methods. *Remote Sensing of Environment*, *113*(6), 1133-1147.

Yu, Q., Gong, P., Clinton, N., Biging, G., Kelly, M., & Schirokauer, D. (2006). Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogrammetric Engineering and Remote Sensing*, *72*(7), 799-811.

Zhang, Y. (2001). Texture-integrated classification of urban treed areas in high-resolution color-infrared imagery. *Photogrammetric Engineering and Remote Sensing*, *67*, 1359-1365.

Zhou, W., Huang, G., Troy, A. and M. Cadenasso. (2009). Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: a comparison study. *Remote Sensing of Environment, 113,* 1769-1777.

Chapter 3. Developing statewide high-resolution land cover maps for assessing tree resources in the central United States using NAIP imagery

Dacia M. Meneguzzo, Greg C. Liknes, Todd A. Kellerman, Darci A. Paull *(draft manuscript, intended outlet: Remote Sensing)*

3.1 Synopsis

High-resolution land cover maps (1-5 m spatial resolution) produced at a statewide scale are generally unavailable, especially in the central United States. Such geospatial datasets are needed by natural resource professionals in this region for decision-making purposes regarding rare but important tree resources. The ≤ 1-meter spatial resolution, repeat coverage, and affordability of digital imagery from the National Agriculture Imagery Program (NAIP) makes it a desirable data source from which to create land cover data products. However, working with this imagery is very challenging because of processing large data collections with variable quality amongst the many images, which is caused by changing atmospheric and illumination conditions encountered during the several month-long acquisition process. This paper describes a semi-automated, operational approach that uses object-based image analysis, supervised classification, and NAIP imagery as the sole data source to create land cover maps for Kansas and Nebraska, USA; maps for North Dakota and South Dakota are forthcoming. The maps were found to accurately represent tree canopies (98.5 and 98.6% accuracies at the state level for Nebraska and Kansas, respectively), thus providing data at an appropriate scale for land managers in this region. The resultant datasets have been published as research datasets that are available for download at no cost.

3.2 Introduction

Tree-covered lands in the Great Plains region of the central United States today consist primarily of trees naturally occurring along stream and river corridors, and

trees established for agroforestry purposes, such as windbreaks. Trees can also be found in small woodlots and scattered across pastures and rangelands. Although scarce in terms of overall areal coverage, these tree resources provide a wide variety of important economic and ecological benefits (Guo et al. 2004), such as: sequestering carbon; enhancing soil, water, and air quality; conserving biodiversity; protecting soil, livestock, crops, and wildlife from harsh winds and snow; mitigating livestock odors; and improving aesthetics (Jose 2009). De Foresta et al. (2013) list timber, fuelwood, marking ownership boundaries, living fences, providing non-traditional crops of fruits and nuts, shade, aesthetics, and odor control as additional functions. Bentrup et al. (2019) also state that trees and shrubs in agroforestry practices provide benefits to pollinators, including habitat, landscape connectivity, and reducing exposure to pesticides.

Carbon sequestration and energy conservation and supply are also key benefits provided by small forest fragments. While forests are often considered and studied as large carbon stocks, trees in agroecosystems are gaining recognition for their potential to store large amounts of carbon and there is interest in quantifying this (Brandle et al. 1992, Kort and Turnock 1999, Montagnini and Nair 2004, Schoeneberger 2009, Czerepowicz 2012). In terms of energy conservation, trees help reduce energy consumption by protecting homes and structures from adverse weather conditions, such as wind and intense sunlight, and provide a source of renewable energy.

While trees provide many important functions in rural landscapes, information about their location, extent, and function is severely lacking. There are two major natural resource inventory systems in the United States: the USDA Natural Resources Conservation Service's National Resources Inventory (NRI) and the Forest Inventory and Analysis (FIA) program conducted by the USDA Forest Service. Both inventories collect information about forest land but employ definitions that have minimum area and width requirements that often exclude trees in agroforestry practices, e.g., windbreaks and riparian buffers, due to their small size and/or narrow shape. Furthermore, these sample-based efforts are

conducted at scales (e.g., one plot per 6,000 acres in the case of FIA) that are too coarse to adequately capture such small tree features within geographic extents of interest to land managers. As such, these trees are referred to as ‘’trees outside forests’’, or TOF, and their extent remains largely unknown.

This data gap is especially problematic in the wake of climate change and the introduction of invasive pests, such as the emerald ash borer (*Agrilus planipennis* Fairmaire). Such pests do not discriminate between trees on definitional forest lands versus those on nonforest lands, so more comprehensive data that includes all tree resources are desperately needed, especially in the Great Plains region where there are many TOF. To address this, the Great Plains Initiative (GPI) was formed via a cooperative project with the USDA Forest Service and state forestry agencies in North Dakota, South Dakota, Nebraska, and Kansas. Their goal was to conduct a ground-based inventory of trees in areas that do not meet FIA’s definition of forest land for their respective states (Lister et al. 2012). The results of this inventory (field data collected in 2008-2009) indicate that the area of these TOF lands was approaching that of FIA definitional forest land for the four-state region, with an estimated 5.1 million acres of TOF lands compared to 6.4 million acres of forest land (2009 estimate) (Meneguzzo et al. 2018).While this effort was valuable in that it provided previously unavailable field-based information about TOF, it was a one-time effort and the plot-based data fail to provide detailed spatial information about TOF. Two additional studies also attempt to quantify the data gap between FIA definitional forest land and TOF: one is a study of “working trees” by Perry et al. (2009) concluded that including non-definitional forest land with trees in the inventory would increase estimates of total canopy-covered land by at least 25% in Kansas, Nebraska (26%), North Dakota (38%), and South Dakota (30%), USA; the other occurred in the neighboring state of Iowa, where an analysis of a 2009 High Resolution Land Cover dataset (Iowa DNR 2017) indicated that area of tree-covered lands was 14% greater than that of FIA forest land (Nelson et al. 2016). Interestingly, this was similar to the result of 19% for Iowa reported by

Perry et al. (2009). While Meneguzzo et al. (2018) reported that approximately 44% of all treed lands in the four-state study area are occupied by TOF, additional TOF in the same geographic region ranged from only 25-38% per state, or about 29% for the four states combined, according to Perry et al. (2009), suggesting some inconsistency among estimates.

A remote sensing-based approach that maps all tree cover could fill this data gap. While existing land cover datasets, such as the National Land Cover Database (NLCD), Cropland Data Layer (CDL) (Wickham et al. 2014), and NLCD Tree Canopy Cover (TCC) (Coulston et al. 2012) provides useful data on forest and tree cover, these products are derived from satellite sensors that are too coarse (i.e., 30m spatial resolution) to accurately depict individual tree crowns, small tree clusters, or trees in narrow, linear or curvilinear configurations that are common in the area of interest. As such, it is recommended the remote sensing-based approaches incorporate sources of imagery with spatial resolutions finer than 5 m (Liknes et al. 2010). Unfortunately, there are no other higher-resolution land cover datasets available for the broad extent under consideration for this study (but see Iowa DNR (2017)) for a description of a related dataset in neighboring Iowa) so we opted to develop an approach to create our own dataset that focus on tree cover. Ultimately, the goal was to produce such datasets for four states (North Dakota, South Dakota, Nebraska, and Kansas), which required development of a semi-automated, operational mapping approach.

We proposed that land cover maps derived from very high resolution imagery (≤ 1 m) would provide natural resource professionals and managers in the Great Plains region with tree cover information at a scale appropriate for their needs. In particular, we chose to use digital aerial imagery from the National Agriculture Imagery Program (NAIP) because of its 1 meter spatial resolution, nationwide coverage on a repeat basis, very low cloud cover, and public availability at no cost. Although NAIP imagery has its advantages, there are many challenges when working with it, especially over large geographic areas such as a state or multi-state region. For example, NAIP data are acquired over

several months under varying atmospheric and illumination conditions, which results in added complexities and spectral inconsistencies in the imagery. Additionally, some areas require new images be acquired so there can be abrupt temporal and spectral differences within adjacent areas due to acquisition dates being several months apart. In fact, the 2014 collection of NAIP imagery for Kansas did not meet quality standards so it was acquired again in 2015. These 2015 images are of superior quality and are used in this study. For more in-depth information on NAIP imagery and the associated challenges when conducting land cover classification over broad spatial extents, the reader is referred to Maxwell et al. (2017 and 2019).

Other considerations for developing this approach include choosing a classification scheme and method, image pre-processing, incorporating other ancillary datasets, assessing classification accuracy, post-processing and data delivery. In addition, when dealing with such a large study area, other factors must be seriously considered, such as data processing, management, storage, and distribution. Costs related to image acquisition, computing hardware and software, and personnel also play an important role. Lastly, the developed method must be robust and transferable to work for a multi-state region.

At the time this research was originally conducted (circa 2015), NAIP had been used in a variety of studies related to land cover mapping (Iowa DNR 2017, Liknes et al. 2010, Platt and Schoennagel 2009, Davies et al. 2010, Hartfield et al. 2011, Moskal et al. 2011, Hayes et al. 2014, Li and Shao 2013, Knight et al. 2013, Meneguzzo et al. 2013, Ghimire et al. 2014, O'Neil-Dunne et al. 2014, Li et al. 2014, Li and Shao 2014, Maxwell et al. 2014, Qiu et al. 2014, Hulet et al. 2014, Basu et al. 2015, O'Neil-Dunne et al. 2014), but the list shrinks when we consider only those conducted at larger geographic scales. One of the earliest examples is a published land cover map with 15 land cover classes for the state of Iowa, entitled "High Resolution Land Cover of Iowa in 2009"; it was produced using a pixel-based classification approach (Iowa DNR 2017). Other published studies with extensive study areas include Pembina County in North Dakota,

USA (2,900 km^2) (Liknes et al. 2010), Steele County in Minnesota, USA (1,110 km^2) (Meneguzzo et al. 2013), Tippecanoe County in Indiana, USA (1,300 km^2) (Li and Shao 2014), and a 14-county study area in Kansas, USA (32,322 km^2) (Ghimire et al. 2014). The University of Vermont Spatial Analysis Laboratory developed an object-based approach that used NAIP and other ancillary data to map urban tree cover for many cities and counties (70+) in the United States and Canada (O'Neil-Dunne et al. 2014); another result of this was a statewide tree canopy map for the state of Maryland (25,640 km^2) (O'Neil-Dunne et al. 2014). However, the largest land cover mapping project based on NAIP, by far, was conducted by Basu et al. (2015) who mapped tree cover for the state of California, an area of 423,970 km2. Since this research was conducted, the list of large-extent NAIP-based studies has grown slightly to include those by St. Peter et al. (2018) who created pixel-level probabilistic land cover maps for a cumulative area of approximately 116,500 km^2 in multiple states in the southeastern U.S., Maxwell et al. (2019) who mapped land cover for the entire state of West Virginia (approximately 62,000 km^2), USA, and the Chesapeake Bay High-Resolution Land Cover Project which mapped land cover for an area nearly 260,000 km^2 in size spanning multiple states (Chesapeake Conservancy 2016).

While the previously mentioned studies use a variety of classification methods, including pixel- and object-based classifiers, we chose to use geographic object-based image analysis (GEOBIA) with a data mining approach. What distinguishes our study from other NAIP-based classification studies is that we did not include any ancillary data in the segmentation or classification processes. We found only two other studies, besides our own previous research (Liknes et al. 2010, Meneguzzo et al. 2013), that used NAIP as the single data source; one study mapped urban vegetation (Li and Shao 2013) while the other mapped pinyon-juniper cover in the western U.S. (Hulet et al. 2014). They successfully accomplished their study objectives using GEOBIA and NAIP. While

encouraging, these studies had different objectives and we still needed to develop our own approach to complete this large-scale mapping task.

The objective of this research is to develop a robust and transferable method for operational mapping of general categories of land cover in rural landscapes using 1 m NAIP imagery as the sole data source for a four-state study area (Kansas, Nebraska, North Dakota, and South Dakota) in the Great Plains region of the central United States. An additional goal was to develop a single segmentation routine that can be applied within the four-state study area without having to make adjustments when moving from county to county or state to state. Building on our prior success mapping tree cover from NAIP imagery (Liknes et al. 2010, Meneguzzo et al. 2013), we pursued the development of an operational method that used a GEOBIA approach with the Random Forests (RF) classifier (Breiman 2001a) to map general categories of land cover. This paper presents an updated method of our most recent research (Meneguzzo et al. 2013) for large-area mapping; improvements made to our previous method include: 1) more semi- and automated processing to make the approach operational, 2) adding more land cover classes, 3) building a segmentation routine that works generally well across a variety of landscapes so settings do not have to be adjusted within the four-state study area, 4) developing a more robust classification approach to accommodate the spectral and illumination issues associated with NAIP imagery. Because of the large size of the study area (nearly 797,000 km^2 when completed), the processes of image segmentation and model development had to take speed and simplicity into consideration; there is a tradeoff between accuracy and efficiency. Previously, we concluded in Meneguzzo et al. (2013) that urban areas require a different segmentation routine and classification model than rural areas due to the increased complexity of those landscapes. As a result, this study will focus on rural lands only.

3.3 Materials and Methods

3.3.1 Study Areas

The states of Kansas (213,096 km^2) and Nebraska (200,519 km^2), USA, lie in the heart of the Great Plains region of the central United States and rank 13th and 15th in state land area, respectively (Figure 3.1). The predominant land covers in both states are related to agricultural land uses. Data obtained from the 2015 cropland data layer (USDA National Agricultural Statistics Service Cropscape) indicate that the top three land cover classes in Nebraska are grassland/pasture (53 percent), corn (19 percent), and soybeans (10 percent). Grassland/pasture is also the most common land cover class in Kansas, making up nearly 46 percent of the total area, followed by winter wheat (17 percent), and corn (8 percent). Deciduous forest makes up nearly two percent of the total land area in Nebraska and 5 percent of Kansas' total area but this does not include TOF or trees in other small and/or narrow configurations. According to the 2017 census of agriculture, land in farms totaled almost 45 and 46 million acres in Nebraska and Kansas, which make up 91 and 87 percent of the state's total areas, respectively. Kansas accounts for five percent of U.S. agriculture sales, while Nebraska has 6 percent (USDA National Agricultural Statistics Service 2017). Both states are among the least densely populated states in the U.S. (Pariona 2017). The population of Kansas is 2,913,314 and ranks 35th while the population of Nebraska is an estimated at 1,934,408 and ranks 37th (Sawe 2017). This equates to population densities of nearly 14 people per square kilometer in Kansas (10th) and about 10 people per square kilometer in Nebraska (8th) according to the 2016 U.S. census bureau estimates (Pariona 2017).

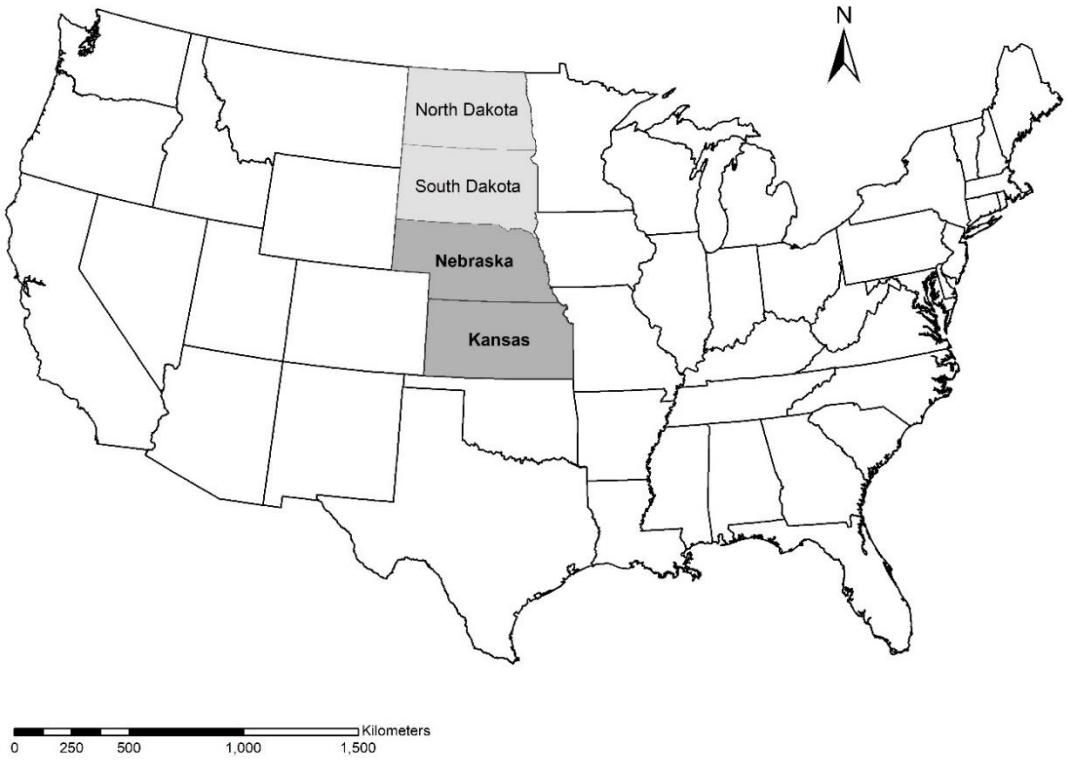


Figure 3.1. Location of Kansas, Nebraska, South Dakota, and North Dakota in the central United States.

The climate of Nebraska is humid in the east and semi-arid in the west with hot and humid summers and cold, harsh winters (Weather Atlas: Nebraska n.d.). Summer daytime high temperatures average 29 to 32°C while winter, i.e., January, average low temperatures range from -12 to -7°C. The statewide average rainfall is 635 mm and decreases, along with humidity, moving from east (813 mm) to west (356 mm). Kansas climate is similar, except for the southeastern portion of the state which has milder winters (Weather Atlas: Kansas n.d.). Both states lie in the “Tornado Alley” of the United States and are subject to other extreme weather events, such as severe thunderstorms, droughts, flooding, and dust storms (Weather Atlas: Nebraska n.d., Weather Atlas: Kansas n.d.).

3.3.2 Data

NAIP imagery was selected for this study because it: 1) is acquired during the summer growing season so it is leaf-on, making it easier to identify tree canopies 2) has a spatial resolution ≤ 1-meter, so individual trees as well as single rows of trees are easily identifiable, 3) includes the near-infrared (NIR) band, which is useful for distinguishing between vegetation and non-vegetation, and 4) is available for all states in the conterminous U.S. at low or no cost to the user. Statewide collections of aerial images were obtained on external hard drives as uncompressed, 4-band (red, green, blue, and NIR) digital ortho quarter quad (DOQQ) image tiles in GeoTIFF format with 300 m of overlap between adjacent images. The Kansas collection is comprised of 5,902 images with acquisition dates between June 20th and October 11th, 2015, while the Nebraska collection contained 5,627 images that were acquired between July 2nd and September 21st, 2014. All told, nearly 2 TB of data required organized data management practices, such as sorting image tiles and segments by county. An alternative would be to use the NAIP compressed county mosaic products but the compression process results in loss of image information, which can lead to reduced classification accuracy (Zabala et al. 2012). As such, we opted to use uncompressed images to retain all spectral information.

Because the focus of this study is on tree resources in rural landscapes, classifying land cover in cities and towns is outside the scope of this study and these areas were masked out and assigned to their own class during a post-processing phase. The U.S. Census Bureau “Incorporated_Place” geospatial layer was used to identify and reclassify the locations of cities and towns. No other datasets were utilized in this study. While light detection and ranging (LiDAR) data may seem like a desirable option to help identify tall tree features in relatively flat topography, LiDAR data were spotty in terms of coverage and collection dates and the disparate datasets were inconsistent in format.

Secondly, the processing, storage, and management of LiDAR data would have exceeded available computing resources and available staff time. Therefore, our goal was to create a simpler approach that was not dependent on LiDAR.

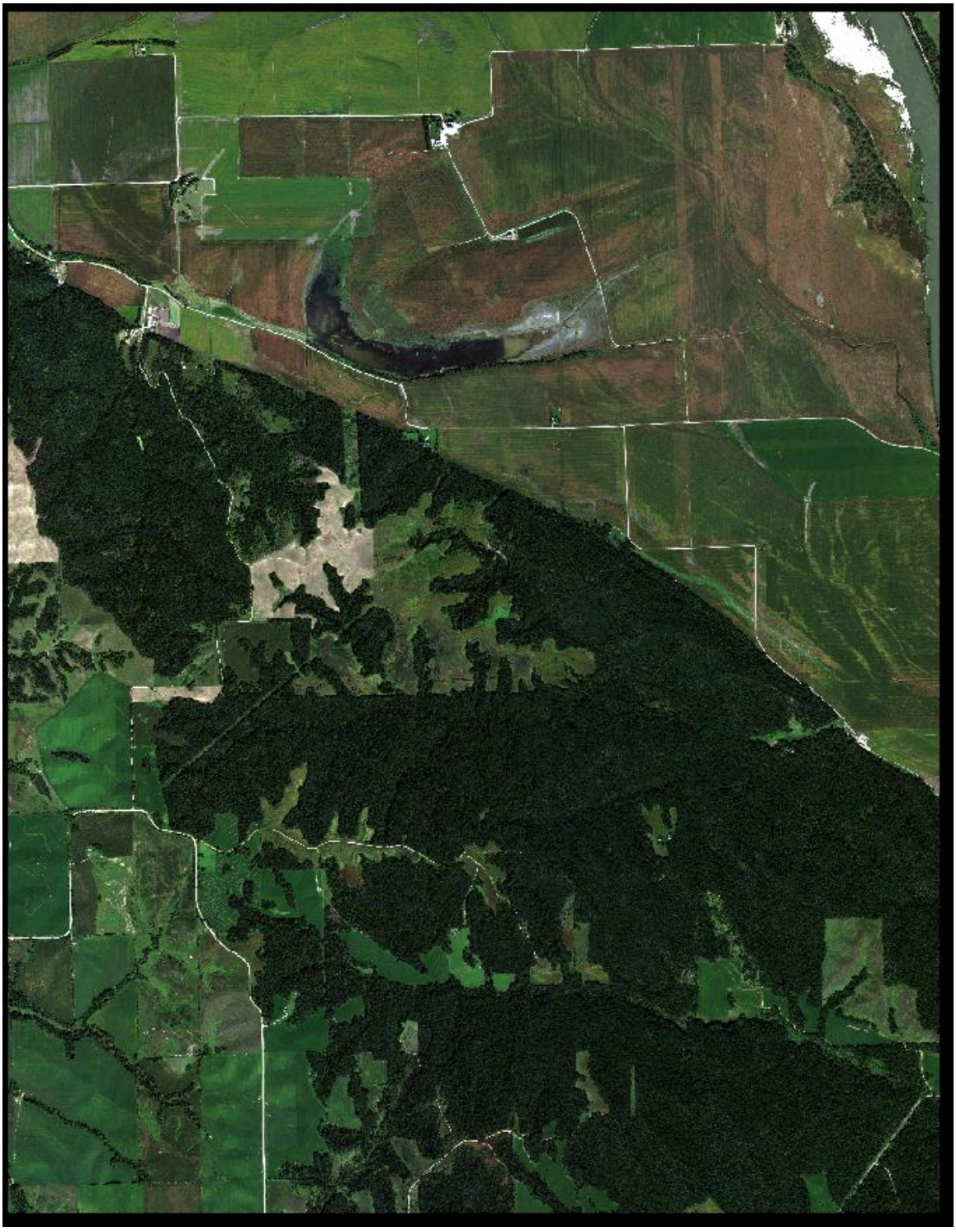
3.3.3 Geographic Object-based Image Analysis (GEOBIA) Approach

3.3.3.1 Segmentation

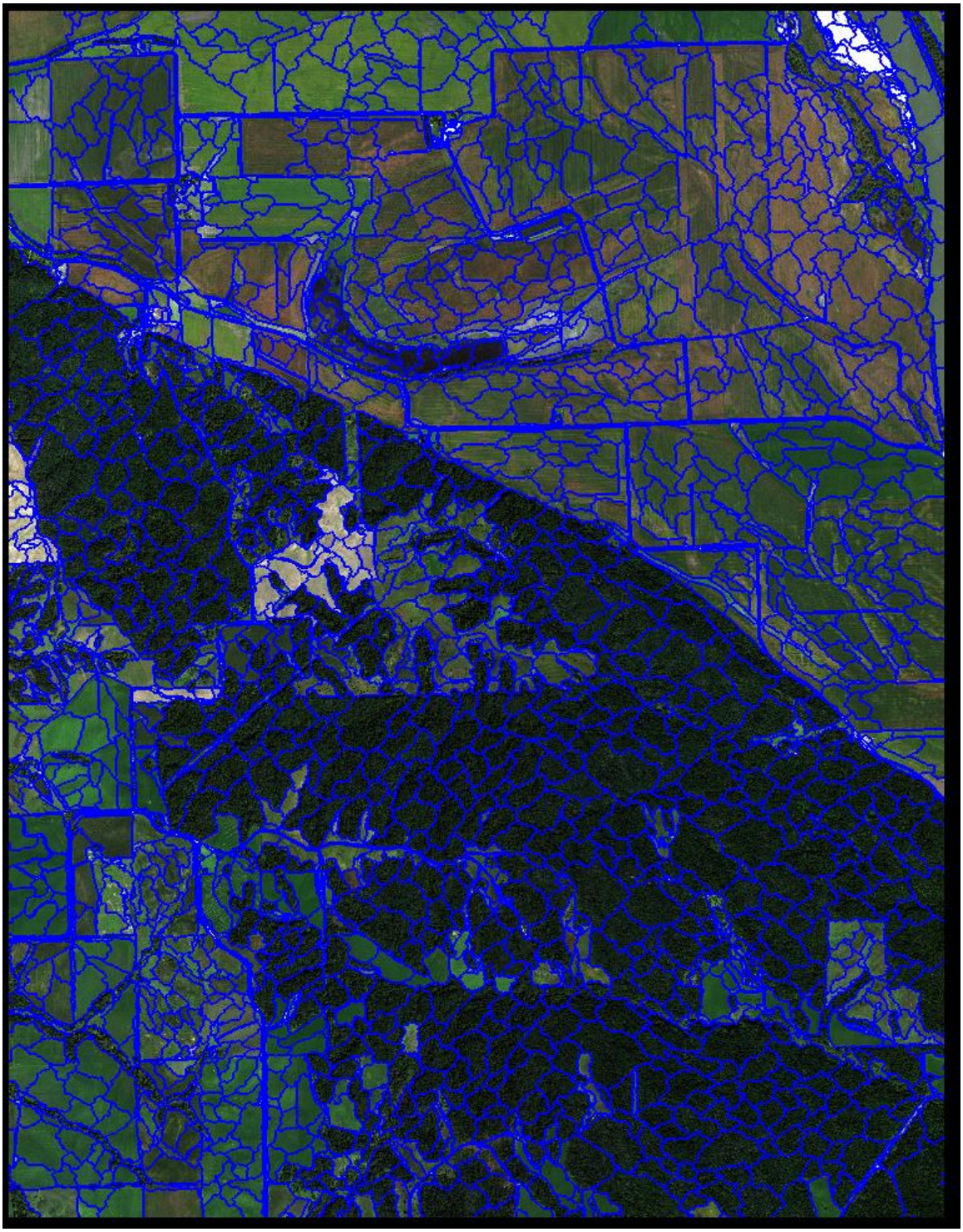
GEOBIA approaches consist of two main phases: image segmentation and classification (Blaschke et al. 2000). Segmentation is a process that divides an image into segments, also called image objects, which represent relatively homogeneous landscape features of interest to be classified (Blaschke 2010); this is an important first step because the segments become the classification units (Benz et al. 2004). The segmentation process for this project was carried out using eCognition® Developer software. While there are many segmentation algorithms available, this project employs the commonly used and very successful 'multiresolution segmentation' algorithm (MRS), which is a Fractal Net Evolution Approach (FNEA) embedded in the eCognition software (Baatz and Schape 2000, Hay et al. 2003, Laliberte et al. 2007, Myint et al. 2011, Witharana and Civco 2014). Starting at the pixel-level, adjacent similar pixels are merged to form larger objects based on homogeneity criteria (Benz et al. 2004, Yu et al. 2006). The formation of image objects is controlled by the scale parameter, which includes assigning weights to shape (color) and compactness, and merging stops when the threshold set by the scale parameter would be exceeded, with a larger scale parameter resulting in larger image objects (Benz et al. 2004, Laliberte et al. 2007). Furthermore, multiresolution segmentation allows the user to extract features at various scales since image objects are formed by user-defined settings that minimize heterogeneity and are not based on the pixel size alone (Benz et al. 2004, Hay et al. 2003). This is an important advantage of GEOBIA because it allows for the creation of image objects at multiple scales, e.g., having objects that represent individual tree crowns as well as larger patches of tree canopy cover within the

same image. We used MRS in our previous tree cover mapping research as well (Liknes et al. 2010, Meneguzzo et al. 2013).

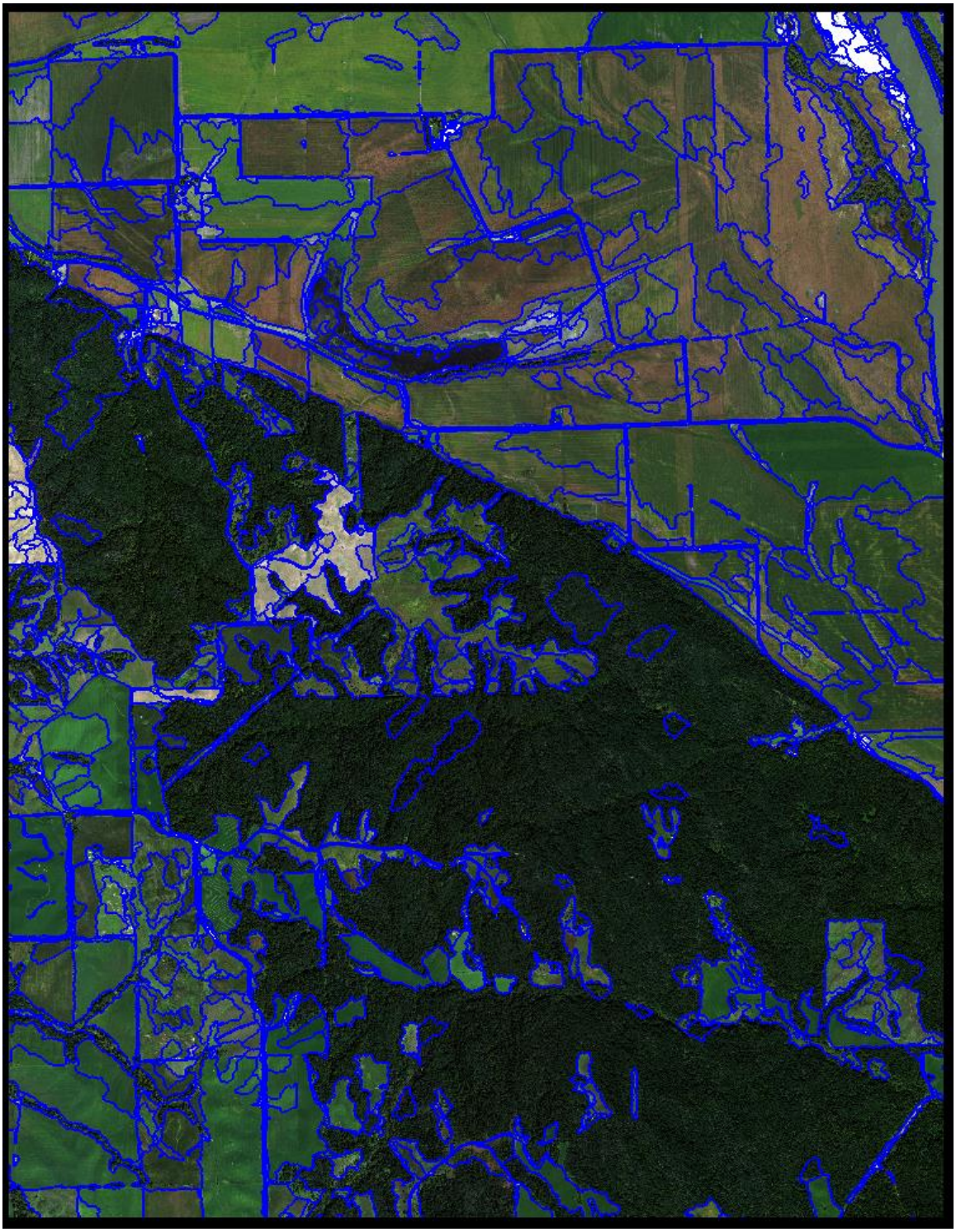
An iterative comparison of potential approaches was used to determine which input layer(s) and segmentation parameters produced the best quality segments to meet our study objectives. We tested numerous combinations of input layers and user-defined settings in various landscapes and then visually inspected the results until we were satisfied with the image objects being created. While our primary focus was to produce segments that accurately represented tree canopy, from single tree crowns to continuous forest canopy, we also wanted segments that depicted other general landscape features, such as roads, farm fields, and wide rivers. We arrived at the following: input layers = green spectral band and an edge-detection layer derived from applying a 3x3 sobel operator filter to the green spectral band (referred to as “sobelop2”); scale parameter = 50; shape = 0.3; compactness = 0.5. A high-pass filter such as the sobel operator is useful for finding edges that define objects of interest in remote sensing studies (Jones and Vaughan 2010). Next, we applied the spectral difference algorithm to the existing results to merge spectrally similar segments into larger, more real world-like image objects, and thus reduce the overall number of segments. This is similar to Hulet et al. (2014) who also used multi-resolution segmentation followed by the spectral difference algorithm to increase median size and simplify image objects. We used the following settings: maximum spectral difference = 5; input layers = red, green, blue, and near-infrared spectral bands all weighted equally. This process is illustrated in Figure 3.2.



(a)



(b)



(c)

Figure 3.2. Example of the image segmentation routine developed for mapping general categories of rural land cover in the central United States. (a) A subset of

an uncompressed NAIP image, acquired on 16 September 2014, in Dakota County, Nebraska, USA, encompassing approximately 46.4 square kilometers. (b) Image objects created using the multiresolution segmentation algorithm in eCognition Developer software. (c) Final image objects after applying the spectral difference segmentation algorithm with a maximum spectral difference of 5 (b).

Next, image objects and 91 attributes associated with each object were exported in vector (shapefile) format. One of the primary advantages of eCognition software is that it can calculate numerous spectral, spatial, contextual, and textural attributes for each image object that can be used as predictor variables in classification model development. Prior to segmentation, several vegetation index layers were created in eCognition and included the normalized difference vegetation index (NDVI), difference vegetation (DV) index, the simple ratio (SR) index, and the green-red vegetation index (GRVI). While these image layers weren't used in the image segmentation process, segment attributes were derived from them because they are useful for distinguishing vegetation from non-vegetation. Texture measures have also been found to be advantageous in land cover classification, especially for distinguishing tree canopy from other land cover categories, including other vegetated classes (O'Neil-Dunne et al. 2014, Myeong et al. 2003, Johansen et al. 2007, Kim et al. 2011). Our previous research (Liknes et al. 2010) had a similar finding in that second-order texture measures based on Haralick et al. (1973) were the most important predictors of tree cover, of which Gray-Level Co-occurrence Matrix (GLCM) homogeneity was the top predictor. The downside of these is that they are time intensive to calculate for individual image objects, thus greatly increasing processing time. Given the thousands of images that needed to be processed and time considerations, we limited the GLCM options to include only GLCM homogeneity and dissimilarity. Spatial attributes, those describing the relationship to neighboring objects, and shape properties were also selected.

Overall, we elected to include a wide range of attributes to take advantage of the range of predictive information available as image object attributes (Table 3.1).

Table 3.1. Image object attributes used as predictor variables in developing land cover classification models for Kansas and Nebraska, USA, using four-band NAIP imagery where band 1= red, band 2 = green, band 3 = blue, and band 4 = near infrared. Vegetation index layers include the normalized difference vegetation index (NDVI), difference vegetation (DV) index, the simple ratio (SR) index, and the green-red vegetation index (GRVI). An edge detection layer derived from applying a 3x3 sobel operator filter to the green spectral band is referred to as “sobelop2”.

Type of feature	Feature from eCognition	Short name
Spectral	Contrast to neighbor pixels (bands 1-4)	CNP_LAYER1, CNP_LAYER2, CNP_LAYER3, CNP_LAYER4
	Contrast to neighbor pixels (selected vegetation index layers)	CNP_NDVI, CNP_GRVI
	Contrast to neighbor pixels (sobelop2)	CNP_sobel
	Mean brightness (average of bands 1-4)	Brightness
	Mean (bands 1-4)	MEAN_LAYER1, MEAN_LAYER2, MEAN_LAYER3, MEAN_LAYER4
	Mean (vegetation index layers)	MEAN_NDVI, MEAN_DV, MEAN_SR, MEAN_GRVI
	Mean (sobelop2)	MEAN_sobel
	Mean of inner border (bands 1-4)	MIB_LAYER1, MIB_LAYER2, MIB_LAYER3, MIB_LAYER4
	Mean of inner border (selected vegetation index layers)	MIB_NDVI, MIB_GRVI
	Mean of inner border (sopelop2)	MIB_sobel
	Mode (median) (bands 1-4)	MM_LAYER1, MM_LAYER2, MM_LAYER3, MM_LAYER4

	Mode (median) (vegetation index layers)	MM_NDVI, MM_DV, MM_SR, MM_GRVI
	Mode (median) (sobelop2)	MM_sobel
	Standard deviation (bands 1-4)	STDDEV_1, STDDEV_2, STDDEV_3, STDDEV_4
	Standard deviation (vegetation index layers)	STDDEV_NDVI, STDDEV_DV, STDDEV_SR, STDDEV_GRVI
	Standard deviation (sobelop2)	STDDEV_S
	Standard deviation to neighbor pixels (bands 1-4)	STDDEVNP_1, STDDEVNP_2, STDDEVNP_3, STDDEVNP_4
	Standard deviation to neighbor pixels (selected vegetation index layers)	STDDEVNPND, STDDEVNPGR
	Standard deviation to neighbor pixels (sobelop2)	STDDEVNP_S
	Skewness (bands 1-4)	Skewness_1, Skewness_2, Skewness_3, Skewness_4
	Skewness (selected vegetation index layers)	Skew_NDVI, Skew_GRVI
	Skewness (sobelop2)	Skew_sobel
	Ratio (bands 1-4)	Ratio_lyr1, Ratio_lyr2, Ratio_lyr3, Ratio_lyr4
	Mean difference to neighbors (bands 1-4)	MDN_LAYER1, MDN_LAYER2, MDN_LAYER3, MDN_LAYER4
	Mean difference to neighbors (selected vegetation index layers)	MDN_NDVI, MDN_GRVI
	Mean difference to neighbors (sobelop2)	MDN_sobel
	Relative border to brighter objects (bands 1-4)	RBBO_LYR1, RBBO_LYR2, RBBO_LYR3,
	Relative border to brighter objects (selected vegetation index layers)	RBBO_LYR4 RBBO_NDVI, RBBO_GRVI
	Relative border to brighter objects (sobelop2)	RBBO_sobel
	Hue (bands 1-3)	HSI_Hue
	Intensity (bands 1-3)	HSI_Int

	Saturation (bands 1-3)	HSI_Sat
Geometry/ Shape	Asymmetry	Asymmetry
	Border index	Border_ind
	Compactness	Compactnes
	Density	Density
	Elliptic fit	Elliptic_F
	Main direction	Main_direc
	Radius of largest enclosed ellipse	RLEE
	Radius of smallest enclosing ellips	RSEE
	Rectangular fit	RectFit
	Roundness	Roundness
	Shape index	Shapeindex
Geometry/ Based on Polygons	Compactness	Compactpol
Texture After Haralick	GLCM homogeneity (all directions)	GLCM_Homog
	GLCM dissimilarity (all directions)	GLCM_Diss

3.3.3.2 Supervised Classification and RF Classification Model Development

The process of classification and creating the final land cover output products were conducted on a county-by-county basis. This approach was taken for a number of reasons: 1) it captures the variations in landscapes and spectral/illumination conditions at a manageable scale, 2) it is a logical way to store, organize, and manage input data and then distribute the output data in smaller file sizes and allowing data consumers to download data for smaller areas of interest, 3) it enhances workflow efficiency by allowing multiple mapping technicians to work on multiple counties at a time at their own pace without impacting the overall workflow.

The first phase of the classification process consisted of collecting training data samples and was conducted using ArcGIS® software. This step was critically important and quality was emphasized over quantity. For each county, a

spatially balanced sample of 10-20 percent of the DOQQ image tiles were selected to account for spectral issues and varying landscapes. From these, representative training samples of each land cover class were selected by a photo interpreter using the corresponding shapefiles that contained the image objects and their associated attributes. Samples were collected for the following land cover classes: tree cover (class 1), other vegetation (class 2), non-vegetation (naturally barren/lacking vegetation or impervious surface; class 3), and water (class 4). To represent variability in spectral and landscape differences within each county, training data were collected from a range of segment sizes, and image textures and colors for each class. The objective was to collect a minimum of 15 samples of each land cover class from each DOQQ shapefile, and a corresponding minimum of 50 samples per class per county; however, this was not always feasible, especially for the water class, which is rare in many portions of our study area. In those cases, as many good, representative samples as possible were collected. When training data collection was completed for each selected shapefile, the polygons (segments) were merged into one shapefile containing all of training data for the county. An example for Dakota County, NE, is shown in Figure 3.3; a total of 308 training data samples (class 1 = 86; class 2 = 103; class 3 = 104; class 4 = 15) were collected from five different NAIP image tiles (approximately 17% of the total number of image tiles that encompass the county) that were selected in a spatially balanced manner throughout the county to represent various landscape and spectral conditions.

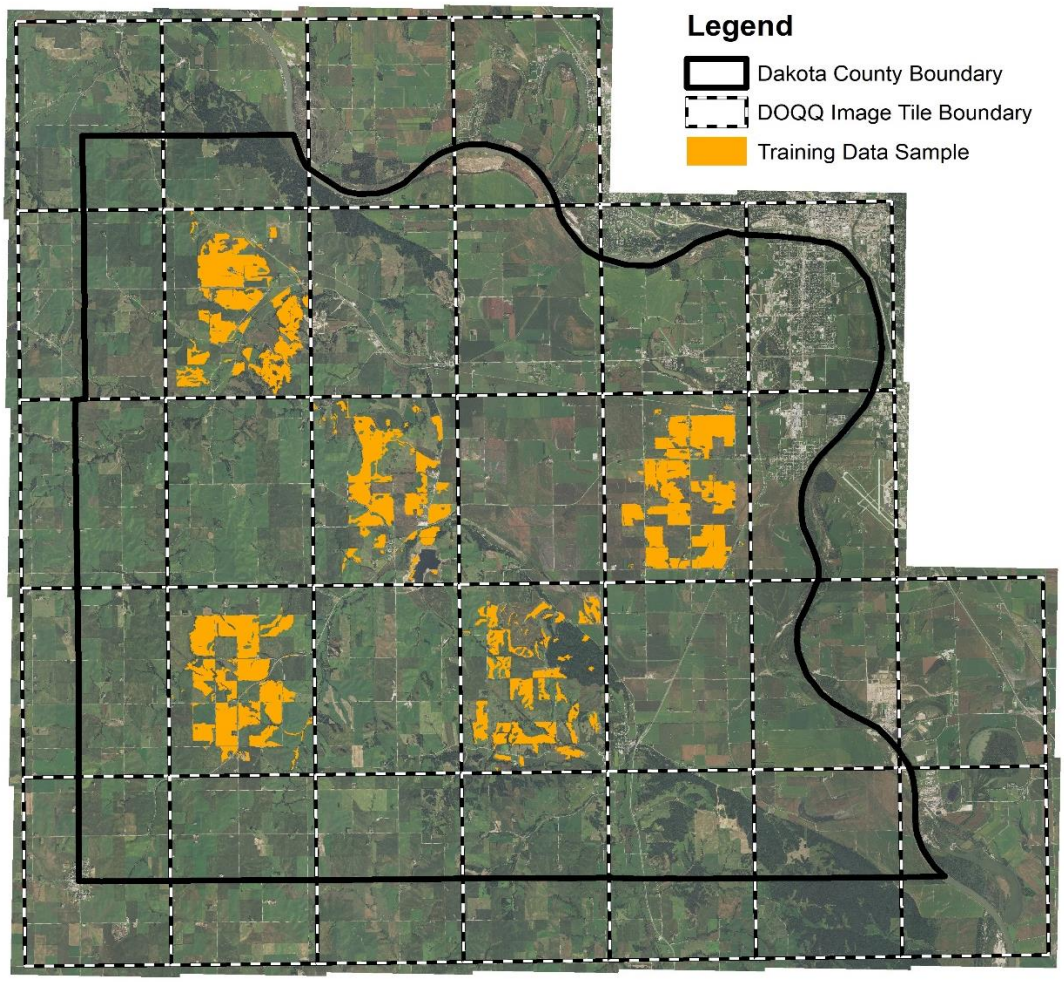


Figure 3.3. Training data collection process for Dakota County, Nebraska, USA. Five shapefiles containing image segments and their associated attributes were selected in spatially balanced manner throughout the county to capture various landscapes and spectral conditions. Training data samples representing four land cover classes were collected from the selected shapefiles.

After the training data were compiled, they were used to develop a classification model using the RF™ algorithm in the freely available R statistical computing environment. The RF model classified the image objects into one of four classes: tree cover, other vegetation, no vegetation (includes naturally

barren areas and impervious surfaces), and water based on its attributes. RF was chosen based on our previous research (Liknes et al. 2010) and the following reasons from (Breiman and Cutler, n.d.): 1) it can accommodate large datasets and thousands of input variables efficiently, 2) it does not overfit the data, 3) input data can be categorical or continuous and no assumptions are made about the distribution of the data, e.g., data do not need to be normally distributed, 4) it produces high classification accuracy, 5) it determines which variables are important in classification and gives them measures of importance, 6) it has ways to balance error in unbalanced datasets, and 7) it provides an unbiased estimate of error based on randomly selected observations that are withheld during model development. Another advantage as noted by Gislason et al. (2006) is that running RF does not require the user to determine variable settings because changing the values of these settings has very little effect on the classification results. Lastly, the r-bridge-install for Python provides a way to connect ArcGIS® and the R computing platform. This allowed us to develop a series of ArcTools that carried out all classification and post-processing steps in ArcMap after collecting the training data, thus making the method more efficient and transferable to mapping partners at other institutions. The ArcTools were assembled into a stand-alone ArcToolbox, which we named the "Plains_Mapping" toolbox and shown in Figure 3.4.

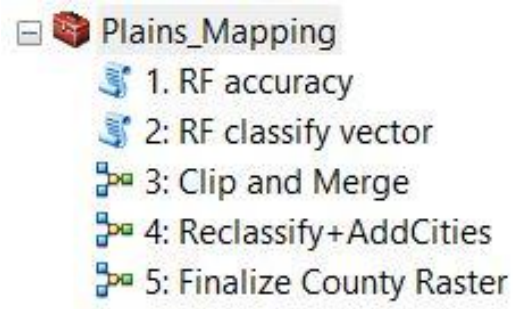


Figure 3.4. Tools in the Plains Mapping ArcToolbox.

3.3.4 Post Processing

After the classification model was applied to all of the shapefiles for the county using the “RF classify vector” tool (tool #2 in Figure 4), a series of post-processing steps were carried out.

1. Clip and merge – each classified shapefile was clipped to the NAIP tile boundary to remove the 300-m overlap areas and then merged into one county-wide shapefile
2. Reclassify and add cities –the “other vegetation” and “no vegetation” classes were combined into one “other land cover” class and the cities and towns were added using the Incorporated_Place layer
3. Manual editing to correct misclassified areas occurred at this point although it is not a tool in the Toolbox
4. Finalize county raster –when editing was completed, this step converted the county-level shapefile to a 4-bit raster with the following land cover class codes: 1 = tree cover; 2 = other land cover; 3 = water; 15 = cities and towns

3.3.5 Assessing Map Accuracy and Area Estimation

To estimate accuracy and area, Olofsson et al. (2014) recommends using reference data that are considered to be of higher quality than the map product, such as field plots, aerial photography, forest inventory data, airborne video, LiDAR, satellite imagery, crowdsourcing, or a reference classification created using a process that is more accurate than the one used to produce the classification under evaluation. In our study we used photo-interpreted segments (described in 3.3.3.2) for assessing accuracy, and FIA field plots for comparing estimates on area (described below).

3.3.5.1 Out-of-bag Error Assessment

As mentioned in section 3.3.3.2, one of the advantages of RF is that it provides an unbiased estimate of error based on observations that are selected at random and withheld during model development; this is known as the out-of-bag (OOB) error that can be used to assess classification accuracy (Breiman 2001a, Breiman and Cutler n.d., Gislason et al. 2006, Cutler et al. 2007). According to Cutler et al. (2007) (p. 2784), “Because the out-of-bag observations were not used in the fitting of the trees, the out-of-bag estimates are essentially cross-validated accuracy estimates.” This information was used to assess accuracy at the county and state levels by calculating class-level producer’s and user’s accuracies as well as overall accuracy from error matrices. At the county level, we created land cover classification models from training data 10 times and averaged the OOB samples in order to produce an estimate of agreement between the training data and the classification model. This process was carried out using the training data collected for each county and the “RF Accuracy” tool as seen in Figure 3.4; the results for each county were recorded and are listed in the supplemental files with the published research datasets (refer to Table 3.2). County-level training data were compiled for each state to build statewide land cover classification models and accuracy was assessed for each state using its OOB sample.

3.3.5.2 Manual Editing and Qualitative Assessment

As mentioned in Section 3.3.4, the classification results were manually edited (after the accuracy assessment) to correct misclassification errors as part of the post-processing procedures. In an additional quality assurance check, several counties were selected in a spatially balanced manner throughout the state and independently reviewed by a third party with GIS and aerial photo interpretation skills. Reviewers were asked to review the output products and make notes of any errors, or repeated types of errors, and provide an overall sense of accuracy.

We recommended the review take place at a scale of 1:15,000. O'Neil-Dunne et al. (2014) states that a scale of 1:3,000 is common but can vary based on map quality, time and labor budget, and preferred accuracy. Given the larger geographic extent of our project area, we felt the 1:15,000 scale was sufficient. Errors observed at a finer scale were not considered significant enough to warrant the time it would take to correct them.

3.3.5.3 Land Cover Area Estimates

Area estimates for the basic land cover categories were obtained from the 1 m statewide land cover classification datasets by summarizing pixel counts of each category. Although Olofsson et al. (2014) recommends producing area estimates from reference data, post-stratified by map class areas, no independent, comparable datasets existed for our multi-state area of interest to serve as such a reference dataset and we did not have the time and/or budget to create one. However, publicly available FIA data are available for land use classes that are similar to our land cover output classes. Using the online EVALIDator estimation tool (https://apps.fs.usda.gov/Evalidator/evalidator.jsp), we obtained estimates of FIA forest land, nonforest land, and water. FIA's formal definition of forest land includes minimum area (0.4 ha) and width requirements (36.6 m), includes areas of regenerating forest having sparse or no tree cover, and excludes tree cover in other land uses (e.g., urban parks). FIA nonforest land class includes all sampled land that does not meet the definition of forest land or water, so it includes cities and towns. While FIA may not be the most ideal reference dataset for this study, we believe it to be of good quality and the best available option for comparison with our map-based estimates.

3.4 Results

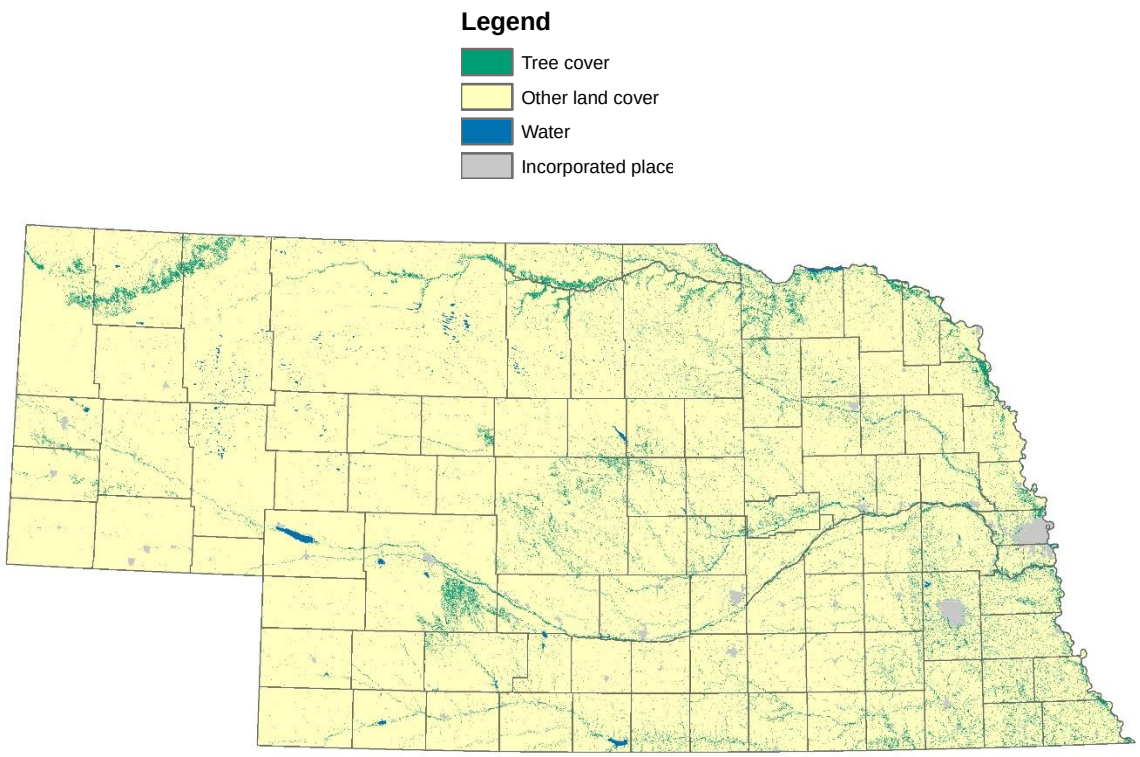
3.4.1 Statewide Datasets

Statewide four-class land cover maps at 1 m spatial resolution have been completed for Kansas (Figure 3.5a) and Nebraska (Figure 3.5b); Figure 3.5(c-d) shows the current mapping status, as of March 2020, for North Dakota and South Dakota, respectively. The high resolution imagery yields spatially explicit detail about tree cover, which can be seen in Figure 3.6.

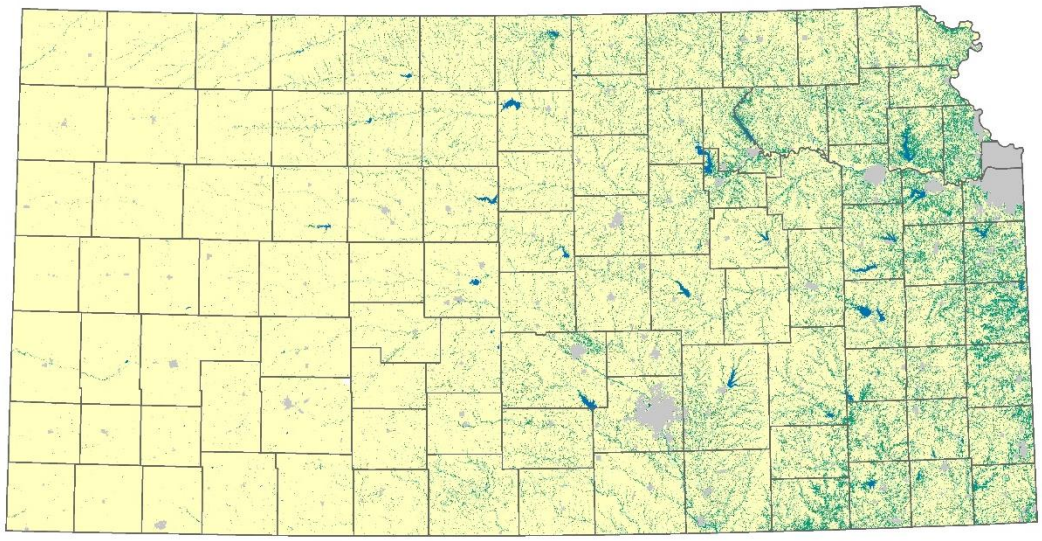
The county-level classification output files are four-bit rasters with the following land cover codes: 1 = tree cover; 2 = other land cover; 3 = water; 15 = cities and towns (from the incorporated places geospatial layer). The land cover maps have been published as research datasets and are freely available in the USDA Forest Service Research Data Archive. Each data publication download contains the county-level output maps, metadata, a file index, and a supplemental file containing county-level accuracy reports from RF (Table 2). The Kansas dataset is called “High-resolution land cover of Kansas (2015)” (Paull et al. 2017) and, similarly, the Nebraska dataset is entitled “High-resolution land cover of Nebraska (2014)” (Kellerman et al. 2019).

Table 3.2. File index for the High-resolution land cover of Nebraska (2014) research dataset publication in the USDA Forest Service Research Data Archive.

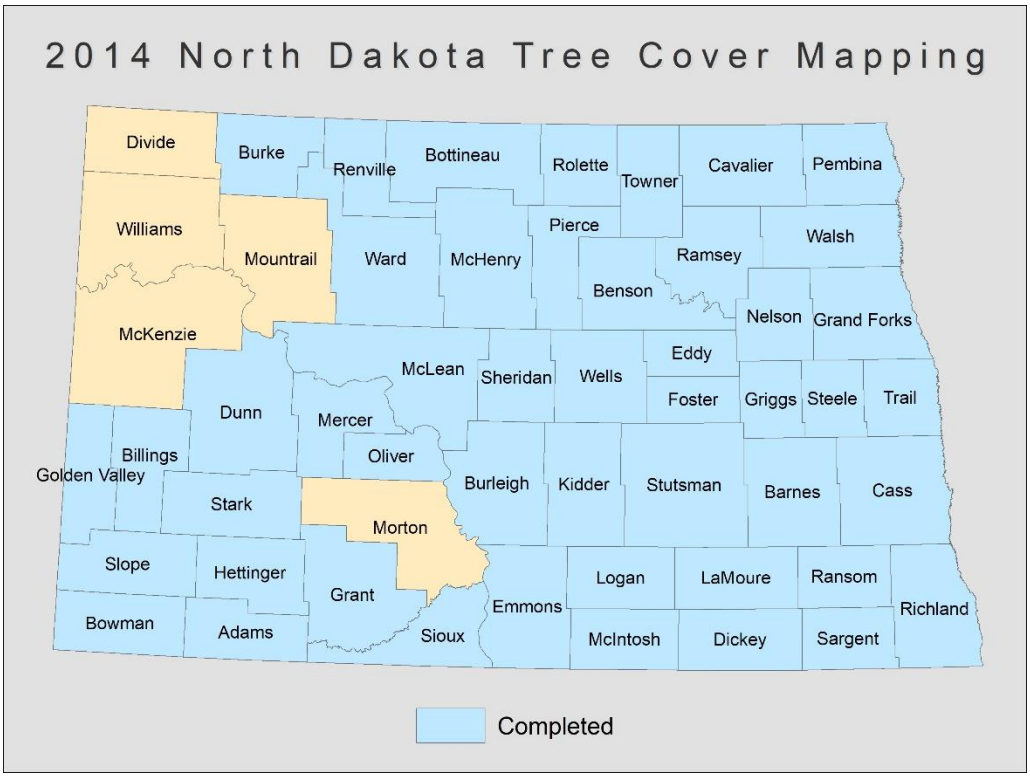
File	Folder	Description
_metadata_RDS-2019-0038.html		Metadata file in HTML format containing a description of the content, quality, and other characteristics of the data.
_metadata_RDS-2019-0038.xml		Metadata file in Extensible Markup Language (XML) format containing a description of the content, quality, and other characteristics of the data.
*COUNTY*_Co.tif	\Data	Georeferenced raster digital TIF files (93) containing 1 meter resolution land cover image for the specified *COUNTY* (and associated files).
*COUNTY*_Co.tif.xml	\Data	Metadata files (93) containing spatial and other information specific to each *COUNTY*, meant to be viewed in conjunction with the associated *.tif file.
NE_2014_county_accuracy_reports.csv	\Supplements	Comma-delimited ASCII text file containing a table showing how well the classification model was able to separate the land cover classes for each county.


Legend
Tree cover
Other land cover
Water
Incorporated place

(a)



(b)


2014 North Dakota Tree Cover Mapping
Divide
Burke
Renville
Bottineau
Rolette
Towner
Cavalier
Pembina
Williams
Mountrail
Ward
McHenry
Pierce
Ramsey
Walsh
Benson
Nelson
Grand Forks
McKenzie
McLean
Sheridan
Wells
Eddy
Foster
Griggs
Steele
Trail
Dunn
Mercer
Oliver
Billings
Golden Valley
Stark
Burleigh
Kidder
Stutsman
Barnes
Cass
Morton
Slope
Hettinger
Grant
Emmons
Logan
LaMoure
Ransom
Richland
Bowman
Adams
Sioux
McIntosh
Dickey
Sargent
Completed

(c)


2014 South Dakota Tree Cover Mapping
Harding
Perkins
Corson
Campbell
McPherson
Brown
Marshall
Roberts
Walworth
Edmunds
Day
Dewey
Ziebach
Potter
Faulk
Spink
Grant
Codington
Butte
Clark
Deuel
Meade
Sully
Hamlin
Hyde
Hand
Lawrence
Stanley
Hughes
Beadle
Kingsbury
Brookings
Haakon
Pennington
Buffalo
Jerauld
Sanborn
Miner
Lake
Moody
Jones
Lyman
Custer
Jackson
Brule
Aurora
Davison
Hanson
McCook
Minnehaha
Mellette
Fall River
Shannon
Tripp
Douglas
Hutchinson
Turner
Bennett
Todd
Gregory
Charles Mix
Lincoln
Bon Homme
Yankton
Clay
Union
Completed

(d)

Figure 3.5. (a) Statewide land cover map of Nebraska, USA, derived from 1 m 2014 NAIP imagery. (b) Statewide land cover map of Kansas, USA, derived from 1 m 2015 NAIP imagery. (c) Statewide land cover mapping project status as of July 2020 for North Dakota, USA (2014 NAIP). (d) Statewide land cover mapping project status as of July 2020 for South Dakota, USA (2014 NAIP).

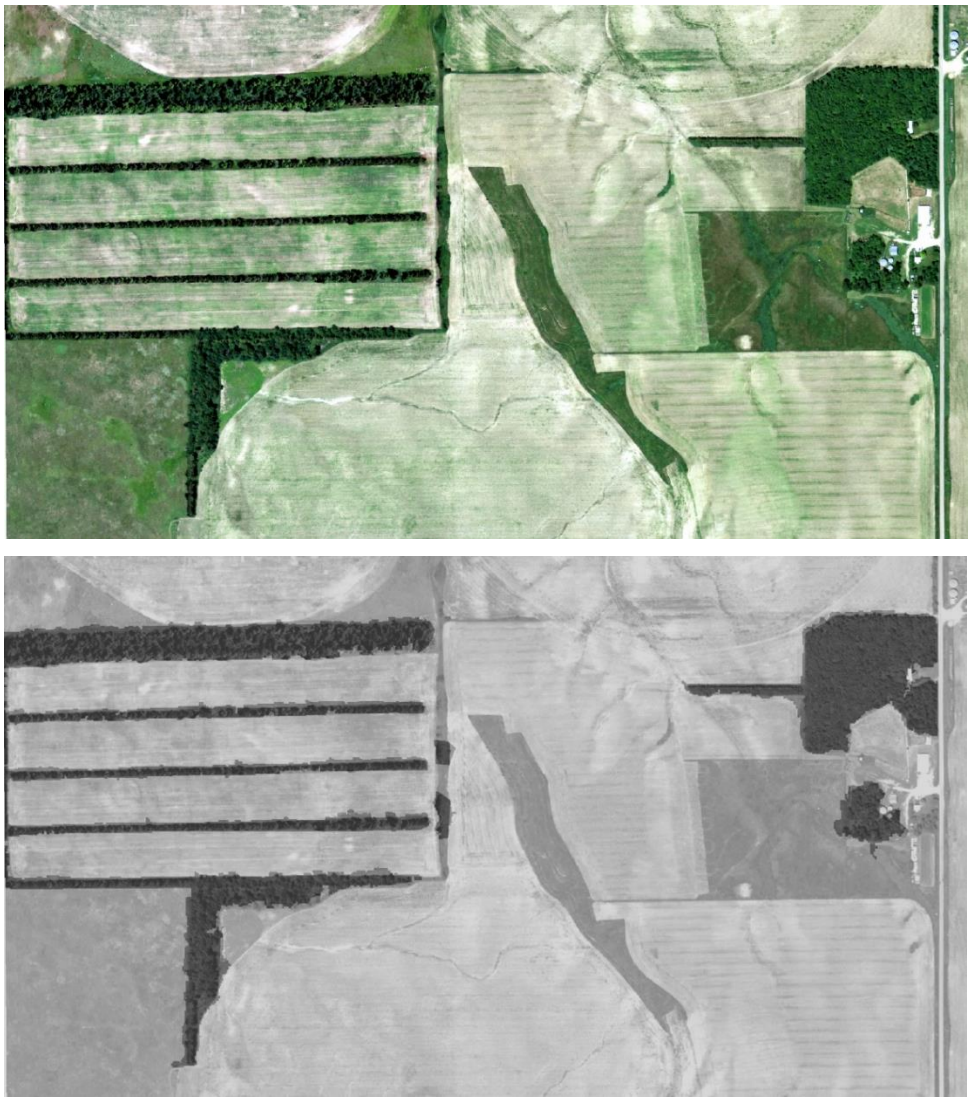


Figure 3.6. A close-up example of uncompressed NAIP imagery (top image) and the resulting 1 m land cover map (bottom image - tree cover shown in dark gray; other land cover in light gray) for an area in rural Antelope County, Nebraska, USA.

3.4.2 Assessing Map Accuracy and Area Estimation

3.4.2.1 Out-of-bag Error Assessment

As described in section 3.3.5.1, the county-level and statewide OOB samples were used to assess accuracy. Collecting training data for each state were labor intensive processes and we didn't have the resources to collect tens of thousands of additional samples to use as validation data in order to conduct another accuracy assessment; we opted to make use of what we had available to us. According to the county accuracy reports (refer to Table 3.2), the total n number of samples for each county ranged from a low of 48 to a high of 930 (average n per county = 436) in Kansas and ranged from 199 to 1,348 (average n per county = 580) in Nebraska. Overall, a combined total of 99,706 photo-interpreted training data samples were observed in Kansas (45,742) and Nebraska (53,964).

We attempted to collect an adequate number of samples for a statistically valid accuracy assessment using the general guideline of at least 50 samples (Olofsson et al. 2014, Congalton and Green 2009) for each of the four land cover classes within each county. While we obtained this goal for the vast majority of the counties in both states, we fell short on some occasions for various reasons, such as few available potential samples and/or human oversight. There were some counties in each state, 12 in Nebraska and 15 in Kansas, that had tree cover n number of samples less than 50, but nearly all counties had at least 30 samples. The water class had a similar issue but this was expected since water can be rare in the semi-arid climate of the Great Plains region. However, this was less of a concern because sampling strategies can be adjusted to focus on class(es) considered to be more important and/or collect fewer samples of classes that show little variability, such as water (Congalton and Green 2009).

Our sample selection scheme was conducted in a spatially balanced manner proportional to the size of the county so that large counties had more samples, thus creating county-level and statewide training datasets that

adequately represent the entire maps for which accuracy is being assessed. Because each map dataset was also visually assessed and manually edited after the accuracy assessment, we believe the final accuracy would be at least as good as the model agreement shown here (Tables 3.3 and 3.4) based on the cross validation of the OOB observations. Generally speaking, the county-level results followed a similar pattern to the statewide results with high class-level and overall accuracies, and the water class having the lowest accuracies.

According to the statewide OOB samples, 44,677 of the 45,742 OOB observations were correctly assigned to one of the four general land cover classes that we intended to model, resulting in an overall accuracy of 97.7% (tree cover class accuracy = 98.6%). In Nebraska, 52,552 were correctly assigned for an overall accuracy of 97.4% (tree cover class accuracy = 98.5%). Tables 3.3 and 3.4 show the statewide accuracy assessment results for Kansas and Nebraska, respectively. Most of the confusion was between classes 2 (other vegetation) and 3 (non-vegetation). We note here that one of the post-processing steps included combining and reclassifying classes 2 and 3 into the “other land cover” class, which increased the overall accuracy to 98.6% in Kansas and 99.0% in Nebraska. This does not affect the accuracy of the tree cover or water classes. We also note that the published county accuracy reports (refer to Table 3.2) report the percent agreement for the published map product classes of tree cover, other land cover, and water classes as well as the overall agreement. The “cities and towns” class is not included because it was determined using a geospatial layer.

Table 3.3. Error matrix for a Random Forests model used to predict general land cover classes in Kansas, USA, based on the out-of-bag sample containing 45, 742 observations.

	Reference					
Classification	**Tree Cover**	**Other Vegetation**	**Non-vegetation**	**Water**	**Row Total**	**User's Accuracy**
Tree Cover	10,751	104	16	23	10,894	98.7%
Other Vegetation	84	12,874	213	4	13,175	97.7%
Non-vegetation	29	212	13,482	90	13,813	97.6%
Water	36	23	231	7,570	7,860	96.3%
Column Total	10,900	13,213	13,942	7,687	45,742	
Producer's Accuracy	98.6%	97.4%	96.7%	98.5%		
Overall Accuracy	**97.7%**					

Table 3.4. Error matrix for a Random Forests model used to predict general land cover classes in Nebraska, USA, based on the out-of-bag sample containing 53,964 observations.

	Reference					
Classification	**Tree Cover**	**Other Vegetation**	**Non-vegetation**	**Water**	**Row Total**	**User's Accuracy**
Tree Cover	11,155	101	17	28	11,301	98.7%
Other Vegetation	98	18,396	445	18	18,957	97.0%
Non-vegetation	33	404	18,077	85	18,599	97.2%
Water	34	29	120	4924	5,107	96.4%
Column Total	11,320	18,930	18,659	5,055	53,964	
Producer's Accuracy	98.5%	97.2%	96.9%	97.4%		
Overall Accuracy	**97.4%**					

Overall agreement for each county was 92.7% or higher in Kansas and 96.5% or higher in Nebraska. When examining the results for our primary class of interest, i.e., tree cover, we find that the class agreement for that class in each county in Kansas was 87.1% or higher and at least 81.1% in Nebraska. However, the majority of counties, by far, had tree cover accuracies greater than 90.0%. In fact, only one county in Kansas and two counties in Nebraska had tree cover accuracies less than 90.0%.

3.4.2.2 Manual Editing and Qualitative Assessment

As mentioned in Section 3.3.4, the classification results were manually edited to correct misclassification errors as part of the post-processing workflow. Similar to O'Neil-Dunne et al. (2014), we found that most time was spent on manual editing. Even recently, Ahles et al. (2016) include making manual corrections as part of their official workflow with the firm belief that it is worthwhile and results in an improved final product. Accuracy assessments were not repeated following manual editing, so we cannot report accuracies for the final map. But as an additional quality assurance check, several counties were selected in a spatially balanced manner throughout the state and independently reviewed by a third party with GIS and aerial photo interpretation skills.

In Kansas, the selected counties included Cherokee, Crawford, Harvey, Kingman, and Sedgwick. Reviewer feedback indicated that the classification output was "very good" and the datasets were accurate representations of land cover in Kansas. The counties selected for review in Nebraska included Dawson, Pierce, and Sheridan. While the classification output was considered good overall, some errors were identified by the reviewers. The main source of the errors were low- or medium-density segments of tree cover that were misclassified as "other vegetation." We believe this was due to the brighter background on hillsides with drier soil conditions showing in the space around

the lower densities of trees within a segment. All noted errors were corrected and we also revisited other similar areas to check for previously missed errors since tree cover was our primary class of interest. Other observed sources of error were shadows misclassified as water and wetland areas with woody vegetation (but not trees) misclassified as tree cover. The combining of classes 2 (vegetation besides tree cover) and 3 (non-vegetation, including impervious surface) alleviated misclassification issues between those classes.

3.4.2.3 Land Cover Area Estimates

Area estimates for the basic land cover categories were obtained from the 1 m statewide datasets. Map-based area estimates from high-resolution land cover datasets indicate that there are 15,439 km^2 of tree cover in Kansas and 8,362 km^2 in Nebraska for a combined total of 23,801 km^2 (Table 3.5). In comparison, the FIA sample-based estimates of forest land area were 10,227 km2 in Kansas (2015) and 6,313 km^2 in Nebraska (2014) for a combined total of 16,540 km^2 of forest land. This is a difference of 7,261 fewer square kilometers compared to the tree cover estimate from the 1 m land cover maps. These results suggest that 1 m land cover maps are filling a data gap by providing spatial information for all tree cover, not just areas that meet the definition of FIA forest land. The map-based area results for the “other land cover” and “cities and towns” classes combined are very similar to FIA’s nonforest land use estimate, which includes urban areas. For example, in Nebraska, summing the map percentages of “other land cover” (94.14%) and “cities and towns” (1.01%) results in a total of 95.15%, which is very similar to FIA’s estimate of 95.93% “nonforest” (Table 3.5). In Kansas, the results for the mapped water class and the FIA estimate of “water” are very close at 1,786 km^2 (0.84%) and 1,885 (0.88%), respectively (Table 3.5). Results for the total land area estimates and percentages of total area were very consistent between the map products and that derived from FIA data for both states.

Table 3.5. Area (km^2) estimates of four land cover classes in statewide land cover maps of Kansas and Nebraska, USA, derived from 1 m NAIP imagery (Map). (Range of per-county percentages shown in parentheses.). The bottom portion of the table contains area (km^2) estimates of three land cover classes from the USDA Forest Service, Forest Inventory and Analysis (FIA) Program based on field- and aerial photo-based information and calculated using the EVALIDator tool, v. 1.8.0.01 (https://apps.fs.usda.gov/Evalidator/evalidator.jsp).

Land Cover Class (Map)	**Kansas Area (km^2)**	**Kansas Area (percentage of total area)**	**Nebraska Area (km^2)**	**Nebraska Area (percentage of total area)**	***Total Area (km^2)***
Tree cover	15,439	7.25 (0.04 – 32.13)	8,362	4.18 (0.14 – 12.77)	*23,801*
Other land cover	191,536	89.89 (1.16* *29.39* – 99.8)	188,552	94.14 (48.09 – 99.55)	*380,088*
Water	1,786	0.84 (0.01 – 5.09)	1,348	0.67 (0.00 – 3.41)	*3,134*
Cities and towns	4,313	2.02 (0.13 – 98.21* *58.72*)	2,019	1.01 (0.01 – 41.33)	*6,332*
Total Area (km^2)	*213,074*	*100.00*	*200,281*	*100.00*	*413,355*
Land Use Class (FIA)					
Forest	10,227	4.80	6,313	3.15	*16,540*
Nonforest	200,985	94.32	192,196	95.93	*393,181*
Water	1,885	0.88	1,835	0.92	*3,720*
Total	*213,097*	*100.00*	*200,344*	*100.00*	*413,441*

*These values are from the same county and represent an extreme case; the values in italics show the next percentage value in the ordered list

When we examine the county-level results for Kansas, we find that Greeley County, which is on the western border of the state, has the smallest area of tree cover with less than 1 km^2 (0.04% of total county area). Linn County, on the eastern border of the state, has the most tree cover with 503 km^2, which is nearly one-third (32.13%) of the county's area. In Nebraska, Grant County, which lies in the west-central part of the state, has the smallest area of tree cover with 2.8 km^2, or just 0.14 percent of county area. Lincoln County lies to the southeast of Grant County and has the most tree-covered area with about 440 km^2, but this is only 6.6 percent of the county's total area because this county is one of the largest in the state. The average per-county percentage of tree cover is 8.3 in Kansas and 4.7 in Nebraska. In terms of area, the average area of tree cover per county in Kansas is 147 km^2 while the county average in Nebraska is about 90 km^2. These county-level estimates are larger than the average forest land area estimates determined by FIA, which are 98 km^2 in Kansas, and 68 km^2 in Nebraska. Interestingly, there were 22 counties in Kansas (Figure 3.7) and 20 in Nebraska (Figure 3.8) where FIA estimated no forest land area, meaning that none of the field-sampled plots occurred on areas that met the FIA definition of forest land. There were no instances of this with the high-resolution land cover data; all counties had tree-covered areas greater than zero.

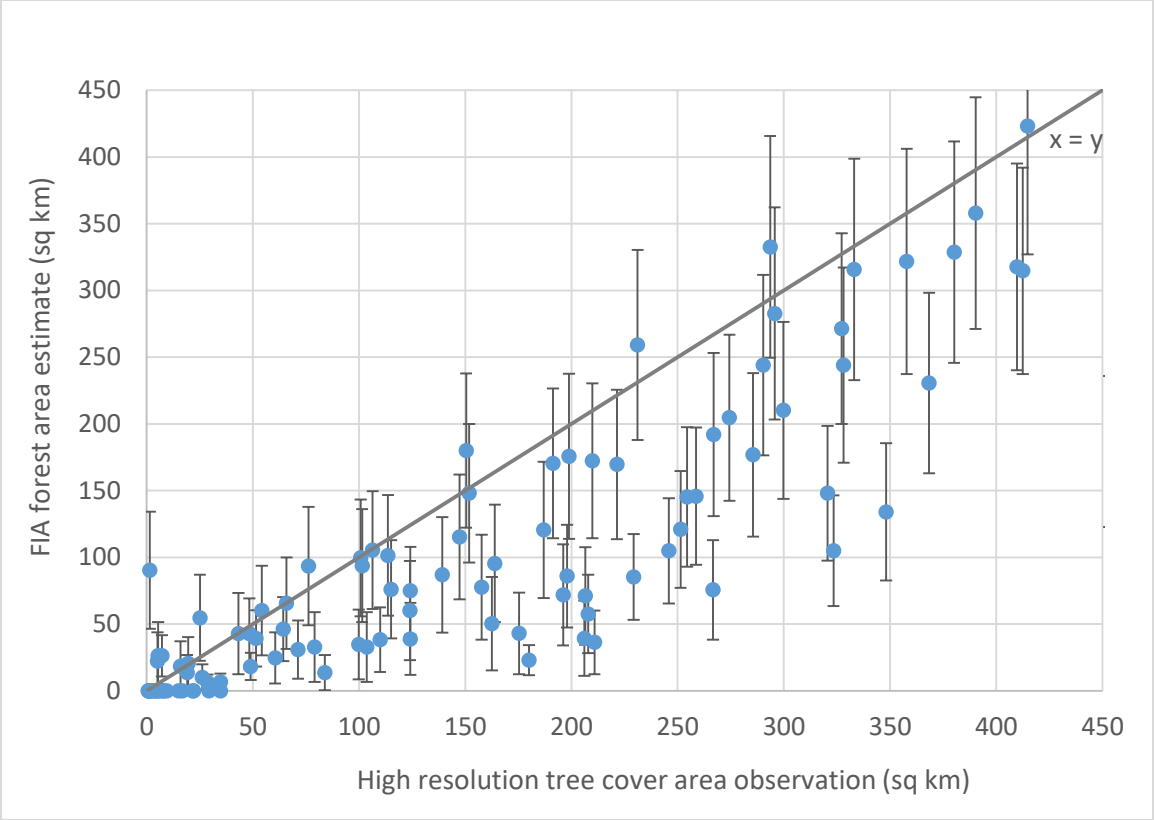


Figure 3.7. Relationship between Kansas per-county estimates of FIA forest land area (2015) and high-resolution (1 m spatial resolution) land cover data (USDA Research Data Archive: https://www.fs.usda.gov/rds/archive/catalog/RDS-2017-0025). Error bars represent the sampling error associated with the FIA estimates at the 68% confidence level.

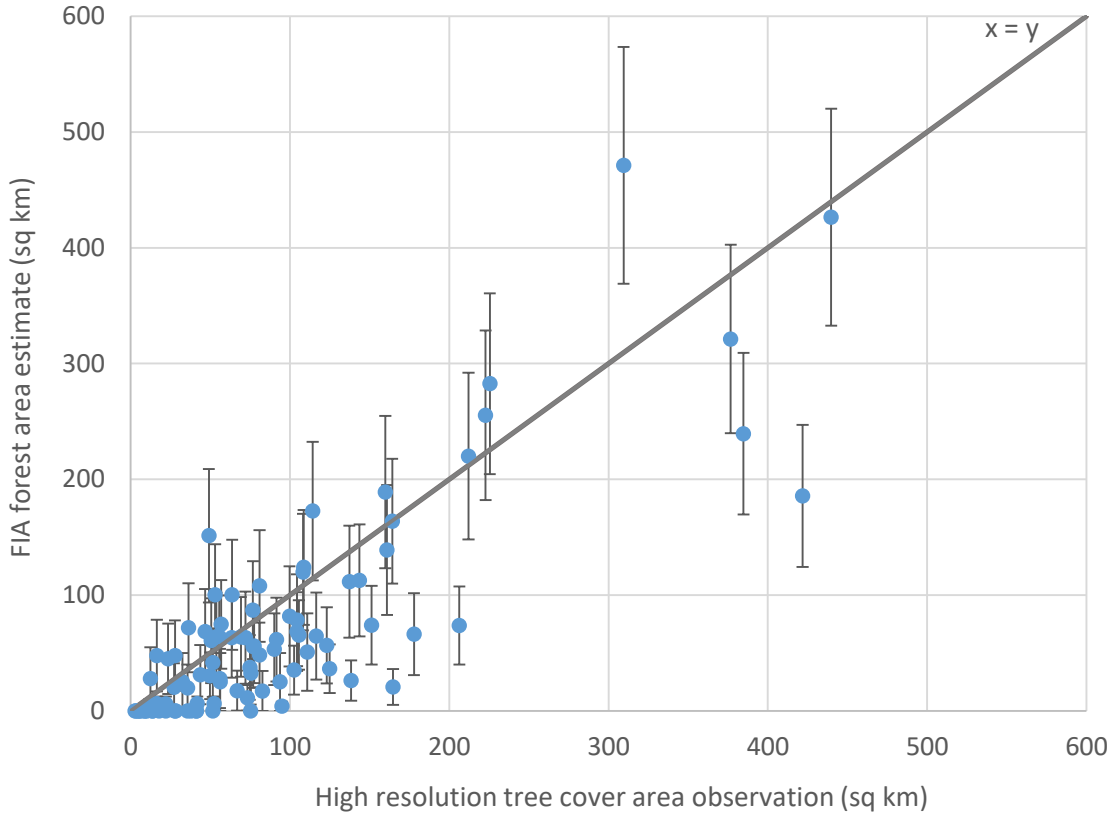


Figure 3.8. Relationship between Nebraska per-county estimates of FIA forest land area (2014) and high-resolution (1 m spatial resolution) land cover data (USDA Research Data Archive: https://www.fs.usda.gov/rds/archive/catalog/RDS-2019-0038). Error bars represent the sampling error associated with the FIA estimates at the 68% confidence level.

3.5 Discussion

The results show that despite the many challenges encountered in working with NAIP imagery, it can be used as the sole data source to produce accurate maps of tree cover and other general land cover classes at a statewide scale in the Great Plains region of the central U.S. Statewide maps have been published for Kansas and Nebraska and maps of North Dakota and South

Dakota are forthcoming. All geospatial products produced as a result of this research are freely available for download. Each county has its own output files so the user can select smaller areas of interest or merge them into one statewide dataset. In addition, the high level of detail in these datasets accurately depicts the small-scale configuration of tree resources in the central U.S. and are, thus, better suited to examine landscape pattern, assess ecosystem services, and assist land managers and decision makers in this region.

The county-level and statewide accuracies as determined by the OOB samples indicate high class-level and overall accuracies. While an additional post-editing accuracy assessment would be ideal, we did not have the data or resources available to collect adequate reference data. However, the OOB accuracy assessment provides sufficient information to evaluate model performance for our purposes. In addition, because each county-level map product was also visually assessed and manually edited after the accuracy assessment, we believe the final accuracy would be similar to the model agreement reported in the results based on the cross validation of the OOB observations.

At the time of this research, there were no standard accuracy assessment procedures for GEOBIA-based maps (O'Neil et al. 2014). Years later, this is still the case (Maxwell et al. 2019, Ye et al. 2018). An extensive review by Ye et al. (2018) examined more than 200 GEOBIA studies and found that many failed to adequately describe their accuracy assessment methods and half used pixel-based rather than object-based approaches to assess accuracy, which presents a mismatch between mapping and assessment unit. Even though half of the studies used a per-object approach, such an approach still has many unresolved issues. They conclude that there is no clear method to assess accuracy of GEOBIA-based maps that have image objects that vary in size and shape. However, we do acknowledge that there are general standards for map accuracy assessment that apply to both pixels and objects, such as properly designing a sampling scheme so reference data samples are selected in an unbiased

manner and ensuring the entire map is represented (e.g., Olofsson et al. 2014, Congalton and Green 2009). We created good training data samples, from which OOB accuracies also were estimated, by attempting to obtain an adequate number of samples, e.g., at least 50, for all land cover classes in each county and samples were collected in a spatially balanced manner proportional to size (i.e., more samples were collected in larger counties) to represent the entire area to be mapped. We acknowledge there is some imbalance amongst the land cover classes in a small number of individual counties due to the scarcity of some classes (i.e., tree cover and water) and/or human oversight during training sample selection. However, all classes had about the same per-class accuracy in all counties, so the disproportional sample sizes are not likely to be a serious issue in this project.

In the future, there are things that could be done to secure a more formal and statistically rigorous accuracy assessment, such as allocating some of the budget up front for designing and conducting the accuracy assessment. Or, we could carefully track editing for a sample of counties to quantify the effect of manual editing to determine when the gains in accuracy offset the time and labor invested in making corrections. Another option would be to prioritize editing to counties where accuracy fell below a particular threshold. This would help increase efficiency for such a laborious process where most time is devoted to editing. Another potential future modification would be to revise the “RF accuracy” tool to issue a warning when sample size is insufficient. And, compiling the training data samples into a statewide dataset gives us a relatively balanced sample with thousands of samples for each land cover class for the statewide accuracy assessments.

Another accuracy assessment challenge we faced was the very large geographic extent of our study area. Even recently Maxwell et al. (2019) (p. 11) states that assessing accuracy was “one of the most challenging aspects” of their study, especially because of their large study area (62,000km^2), i.e., the state of West Virginia, which is less than one-third the size of Kansas alone. Statewide

land cover mapping projects and thus, statewide accuracy assessments, are a rarity. For example, a thorough review study by Ma et al. (2017) examined 173 GEOBIA studies that used supervised classification to map various types of land cover and found that study area size was less than 3 km^2 in almost 96% of the studies; this is 0.001 percent of the size of Kansas. This is one of the strengths of this study; in addition to per-county assessments of accuracy, we were able to obtain statewide assessments for Kansas and Nebraska using the individual statewide OOB samples that, together, included nearly 100,000 photo-interpreted samples collected in a spatially balanced manner over 413,615 km^2.

O'Neil-Dunne et al. (2014) acknowledged that they limited costs by not performing accuracy assessments for all tree canopy mapping projects, particularly when initial accuracy assessments for selected large projects using the same mapping approach were high (90% or better). They also point out that time and money are better spent on making manual corrections that result in higher quality maps with more realistic looking tree canopy while also ensuring that any major errors are avoided; this is reiterated in a more recent paper [68] where making manual corrections is included as part of their official GEOBIA mapping workflow. Lastly, they state that formal accuracy assessments are less of a priority when projects are short on time and budgets, especially considering the lack of standard protocols for assessing accuracy when the units are image objects versus pixels. Based on our similar experience with operational land cover mapping over large geographic areas, we found the above statements to be true. As such, we took a similar approach by prioritizing making manual corrections over trying to implement an additional formal accuracy assessment after the editing phase. While our accuracy results were high, we opted to conduct manual editing to obtain additional improvements (e.g., O'Neil-Dunne et al. 2014). Tree cover is a rare resource in the Great Plains region, and manual editing is a way to ensure that we captured it to the best of our ability but we acknowledge that our product is not perfect.

Classification results also are affected by the quality of the segments. If segments do not accurately represent a class of interest, e.g., by including pixels of another class, the resulting segments will containing a mix of classes; the segment will likely be assigned to the majority class and part of the image object will be misclassified. We did not attempt to alter such segments due to time-limitations and to reduce the risk of introducing topology errors, such as unclosed or sliver polygons as a result of cutting or altering the polygon boundaries. Furthermore, visual assessment indicates that this did not appear to be a frequent problem.

One reason for the high overall accuracy is the small number of land cover classes (Ma et al. 2017). Initially, our goal was to include more land cover classes for future context analyses but during early research, we found that having additional classes meant more complicated segmentation routines and making adjustments as well as dedicating more time to making manual corrections. Ultimately, the need to reduce the editing workload led to the decision to merge classes 2 and 3 into an “other land cover” class. We deemed this appropriate since our primary focus was on the tree cover class and we did not have the personnel resources to edit additional land cover classes over such a vast area. Other data layers (even those with coarser spatial resolution) can be used to provide context information, e.g., the Cropland Data Layer is a good example considering the location of our study area.

Map-based estimates of tree cover were consistently larger than estimates of forest land area produced by FIA’s field sample-based inventory. This is expected because the FIA definition of forest land excludes “trees outside forests” that can be detected in the high-resolution imagery. Subtracting the FIA estimate of forest land from the map-based estimates of tree cover can approximate the area of TOF and compare our findings to those reported by other studies (e.g., Meneguzzo et al. (2018) and Perry et al. (2009)) that quantified the gap between FIA forest land and TOF. We note that while this simple subtraction method does not exactly meet the definition for TOF because

some FIA forest land has no tree cover (e.g., recently disturbed areas that have not yet regenerated), it still provides a recent and reasonable approximation. We report these results in terms of percentage rather than areal estimates because the studies were conducted at differing time periods and the estimates of forest land area have changed over time. FIA estimates were produced to coincide with the years in which TOF data were obtained or created.

Our map-based results indicate that TOF occupy 34% and 25% of all treed lands in Kansas and Nebraska, respectively, and most closely matched those of Perry et al. (2009) who reported that 25% of total treed land in Kansas (26% in Nebraska) was occupied by TOF. In a separate analysis for Iowa that also used high-resolution map-based data (see Nelson et al. 2016), it was found that TOF made up 14% of all treed lands, which was relatively close to the result of 19% reported by Perry et al. (2009). A similar analysis using the GPI estimates of TOF from Meneguzzo et al. (2018) and FIA forest land estimates produced higher results of 48% and 47% for Kansas and Nebraska, respectively. All studies showed that TOF adds substantial area beyond the area of forest land, but the magnitude of additional TOF area varied among studies. Interestingly, the high-resolution map-based results for Kansas, Nebraska, and Iowa agreed most closely with findings by Perry et al. (2009) even though a different method was used to create the Iowa land cover map. These results demonstrate that 1 m datasets provide more comprehensive information about tree cover, not only in terms of total extent but detailed spatial information as well (see Figures 3.6-3.8).

Overall, our findings agree with the general consensus that land cover mapping using NAIP imagery is challenging. Besides the issue of varying image quality, there are challenges of large amounts of data to manage; segmenting and classifying occur on a per-image tile basis because it’s not possible to mosaic the numerous tiles that make up a statewide collection of image tiles (Maxwell et al. 2019). In addition, we did not have access to eCognition server so segmenting and exporting the attributes took up to six weeks for an individual state; we recommend using eCognition server when available. Batch processing

and Python scripts were used wherever possible to increase efficiency, e.g., sorting the NAIP image tiles by county. While the overall process seems relatively simple, there were a lot of nuances to work through and decisions to be made because of the sheer volume of data we are dealing with. For example, while vector output file format is desirable from an analytical standpoint, we opted to include a post-processing tool that converted the final shapefiles to four-bit raster files in order to keep file size down and increase drawing speed in ArcGIS®.

Future statewide mapping applications may benefit from technological advances such as the rise of cloud computing platforms like Google Earth Engine (GEE), which most likely will alleviate the processing challenges we encountered. In addition, LiDAR data may become a more feasible enhancement to mapping when datasets become more abundant and data processing and storage capacities become more robust. Another step to make the method more efficient would be to reduce the number of input predictor variables by doing more preprocessing, and to exclude or make more efficient the GLCM measures that are so time intensive to calculate. Lastly, we would most likely simplify the classification scheme to create binary maps of tree/not tree cover.

3.6 Conclusions

We proposed and demonstrated how statewide land cover maps for the central U.S. could be derived from 1 m, 4-band (color-infrared) uncompressed NAIP images using GEOBIA and supervised classification with the RF classifier.

We successfully developed one segmentation routine that worked generally well across the various landscapes encountered in the central U.S. It accurately delineated land cover features of interest, especially tree canopies, for rural landscapes in a four-state region that included North Dakota, South Dakota, Nebraska, and Kansas. Adjustments may be needed when applying the approach to other geographic extents beyond our study area. The multiresolution segmentation algorithm utilized information in the green spectral band and one of

its derivatives: an edge detection layer created by applying a 3x3 sobel operator filter to form the image objects. This is different from other statewide land cover mapping studies that used all four spectral bands of NAIP imagery weighted equally for segmentation (Maxwell et al. 2019, Basu et al. 2015).

Implementing this approach on a county-by-county basis was the most practical way to accomplish such a daunting mapping project, and worked across counties of varying size. It is a logical way to organize, manage, and distribute the data. During the training data collection process, it allows the mapping technicians to account for the spectral variability that occurs in the imagery and the changing landscapes within the county by selecting training data samples that represent these varying conditions. And, the creation of a Plains Mapping ArcToolbox made our method available and transferable to our mapping partners. Most natural resource agencies and/or Universities have access to ArcGIS® and the R statistical software is available to everyone for free. The r-bridge install for Python makes it possible to link ArcMap and the RF classifier to conduct the classification process, accuracy assessment, and the post processing steps. In addition, mapping personnel can work on their own counties simultaneously without duplicating efforts.

While this approach was successful, it has its limitations, which we describe in two general categories: human, and computing hardware/software. In terms of the human element, the primary limiting factor in this approach is having a sufficient number of people with the available time and technical skills who are committed to completing the project. At least three full-time staff with entry-level GIS skills and some experience with eCognition Developer software who can commit to a year of time is recommended. This method requires substantial manual interpretation of training data and editing as part of the post-processing procedures. Mistakes as a result of fatigue can be problematic if mapping personnel are limited and the brunt of the work falls on one or two individuals; furthermore, staff turnover, work ethic, and even simple oversights can also have an impact on the results obtained. Recommendations include having an

organized file management system and clear workflow, conduct editing at a scale of 1:15,000 so staff aren't overwhelmed by small details and perfectionism, and administering regular data quality checks to find errors early on.

Computing requirements for both hardware and software are substantial, but careful planning and creating partnerships can limit costs to a manageable level. For example, our mapping partners with the Kansas State University, Kansas Forest Service spent less than $5,000 on building a computer that could quickly and easily handle all associated computing processes, from image segmentation and exporting the attributes to classification model application and post-processing. Although somewhat dated now, we provided them with the following recommendations that can be considered minimal computing specifications for producing similar data products: 64-bit edition of whichever Windows version is purchased; two or more quad core processors or one or more octa core processors; a minimum of 32 GB of RAM; two additional data drives of 4 TB or more. However, not all agencies have equal monetary, computing, or personnel resources, and may not have access to eCognition. In those cases, we (USDA Forest Service) provided the image segments and associated attributes in shapefile format, along with the Plains Mapping ArcToolbox and training manual, and partners were able to successfully complete the mapping classification process without eCognition.

To the authors' knowledge, this is the largest 1 m NAIP-based mapping projects to date. When fully complete with the addition of maps for North Dakota and South Dakota, nearly 797,000 square kilometers will be mapped at 1 m resolution; this is almost 797 billion pixels! This process will include 23,233 image tiles and more than 200,000 files (including intermediate files) totaling more than 4 TB of data. This endeavor provides new data and information about tree resources across the Great Plains regions of the central U.S., at a finer spatial resolution than previously available. Natural resource managers now have the data needed for inventory, monitoring, and decision making at a range of scales:

from individual landholdings to statewide, across the Great Plains region from North Dakota to Kansas.

3.7 References

Ahles, N., MacFaden, S., O'Neil-Dunne, J., Royar, A., Engel, T. GEOBIA systems for massive data processing. 2016. Available online: https://doi.org/10.3990/2.457 (accessed on 17 November 2019)

Baatz, M., Schape, M. 2000. Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation. In Angewandte geographische Informationsverarbeitung XII, Strobl, J., Blaschke, T., Griesebner, G., Eds., Wichmann, Karlsruhe, pp. 12-23.

Basu, S., Ganguly, S., Nemani, R.R., Mukhopadhyay, S., Zhang, G., Milesi, C., Michaelis, A., Votava, P., Dubayah, R., Duncanson, L., et al. A semiautomated probabilistic framework for tree-cover delineation from 1-m NAIP imagery using a high-performance computing architecture. IEEE Trans. Geosci. Remote Sens. 2015, 53, 5690–5708.

Bentrup, G., Hopwood, J., Adamson, N., Vaughan, M. Temperate Agroforestry Systems and Insect Pollinators: A Review. Forests 2019, 10, 981.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS J. Photogramm. Remote Sens. 2004, 58, 239-258.

Blaschke, T. Object based image analysis for remote sensing. ISPRS J. Photogramm. Remote Sens. 2010, 65, 2–16.

Blaschke, T., Lang, S., Lorup, E., Strobl, J., Zeil, P. Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. In Environmental Information for Planning, Politics and the Public, Cremers, A.B., Greve, K., Eds., Metropolis: Marburg, Germany, 2000, pp. 555-570.

Brandle, J.R., Wardle, T.D., Bratton G.F. Opportunities to increase tree planting in shelterbelts and the potential impacts on carbon storage and conservation. In Forests and Global Change: Opportunities for Increasing Forest Cover, Sampson, R.N., Hair, D., Eds., American Forests: Washington D.C., USA, pp. 157-176.

Breiman, L. Random forests. Mach. Learn. 2001a, 45, 5-32.

Breiman, L., Cutler, A. Random Forests. Available online: https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm#papers (accessed on 10 March 2020).

Congalton, R.G., Green, K. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, 2nd ed., Taylor & Francis Group, LLC: Boca Raton, FL, USA, 2009, pp. 63-83.

Coulston, J. W., Moisen, G.G., Wilson, B.T., Finco, M.V., Cohen, W.B., Brewer, C.K. Modeling percent tree canopy cover: a pilot study. Photogramm. Eng. Remote Sens. 2012, 78, 715–727.

Cutler, D., Edwards, Jr., T., Beard, K., Cutler, A., Hess, K., Gibson, J., Lawler, J. Random forests for classification in ecology. Ecology, 2007, 88, 2783-2792.

Czerepowicz, L., Case, B.S., Doscher, C. Using satellite image data to estimate aboveground shelterbelt carbon stocks across an agricultural landscape. Agric., Ecosyst. Environ. 2012, 156, 142-150.

Davies, K.W., Petersen, S.L., Johnson, D.D., Bracken Davis, D., Madsen, M.D., Zvirzdin, D.L., Bates, J.D. Estimating juniper cover from National Agriculture Imagery Program (NAIP) imagery and evaluating relationships between potential cover and environmental variables. Rangel. Ecol. Manage. 2010, 63, 630–637.

de Foresta, H., Somarriba, E., Temu, A., Boulanger, D., Feuilly, H., Gauthier, M. Towards the Assessment of Trees Outside Forests. Available online: http://www.fao.org/3/aq071e/aq071e00.pdf (accessed on 4 December 2019).

Ghimire, K., Dulin, M.W., Atchison, R.L., Goodin, D.G., Hutchinson, J.M.S. Identification of windbreaks in Kansas using object-based image analysis, GIS techniques and field survey. Agroforest. Syst. 2014, 88, 865-875.

Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R. Random Forests for landcover classification. Pattern Recognit. Remote Sens. 2006, 27, 294-300.

Guo, Q., Brandle, J.M., Schoeneberger, M., Buettner, D. Simulating the dynamics of linear forests in Great Plains agroecosystems under changing climates. Can. J. For. Res. 2004, 34, 2564-2572.

Haralick, R., Shanmugan, K., Dinstein, I. Textural features for image classification. IEEE Trans. Syst. Man Cybern. 1973, 3, 610-621.

Hartfield, K.A., Landau, K.I., van Leeuwen, W.J.D. Fusion of High Resolution Aerial Mutlispectral and LiDAR Data: Land Cover in the Context of Urban Mosquito Habitat. Remote Sens. 2011, 3, 2364-2383.

Hay, G., Blaschke, T., Marceau, D., Bouchard, A. A comparison of three image-object methods for the multiscale analysis of landscape structure. ISPRS J. Photogramm. Remote Sens. 2003, 57, 327-345.

Hayes, M.M., Miller, S.N., Murphy, M.A. High-resolution landcover classification using Random Forest. Remote Sens. Lett. 2014, 5, 112-121.

Hulet, A., Roundy, B.A., Petersen, S.L., Bunting, S.C., Jensen, R.R., Roundy, D.B. Utilizing National Agriculture Imagery Program Data to Estimate Tree Cover and Biomass of Piñon and Juniper Woodlands. Rangeland Ecol. Manage. 2014, 67, 563-572.

Iowa DNR. High Resolution Land Cover of Iowa in 2009. Release date 2017. Available online: https://geodata.iowa.gov/dataset/high-resolution-land-cover-iowa-2009 (accessed on 24 March 2020).

Jones, H.G., Vaughan, R.A. Remote sensing of vegetation: principles, techniques, and applications, 1st ed., Oxford University Press Inc.: New York, USA, 2010, p. 146

Johansen, K., Coops, N.C., Gergel, S.E., Stange, Y. Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification. Remote Sens. Environ. 2007, 110, 29-44.

Jose, S. Agroforestry for ecosystem services and environmental benefits: an overview. Agroforest. Syst. 2009, 76, 1-10.

Kellerman, T.A., Meneguzzo, D.M., Vaitkus, M., White, M., Ossell, R., Sorsen, N., Stannard, J., Gift, T., Cox, J., Liknes, G.C. High-resolution land cover of Nebraska (2014). Fort Collins, CO: Forest Service Research Data Archive. 2019. https://doi.org/10.2737/RDS-2019-0038

Kim, M., Warner, T.A., Madden, M., Atkinson, D.S. Multi-scale GEOBIA with very high spatial resolution digital aerial imagery: scale, texture and image objects. Int. J. Remote Sens. 2011, 32, 2825-2850.

Knight, J.F., Tolcser, B.P., Corcoran, J.M., Rampi, L.P. The Effects of Data Selection and Thematic Detail on the Accuracy of High Spatial Resolution Wetland Classifications. Photogramm. Eng. Remote Sens. 2013, 79, 613-623.

Kort, J., Turnock, R. Carbon reservoir and biomass in Canadian prairie shelterbelts. Agroforest. Syst. 1999, 44, 175-186.

Laliberte, A. S., Fredrickson, E. L., Rango, A. Combining Decision Trees with Hierarchical Object-oriented Image Analysis for Mapping Arid Rangelands. Photogramm. Eng. Remote Sens. 2007, 73, 197-207.

Land Cover Data Project. Available online: https://chesapeakeconservancy.org/conservation-innovation-center-2/high-resolution-data/land-cover-data-project/ (accessed on 9 March 2020).

Li, X., Myint, S.W., Zhang, Y., Galletti, C., Zhang, X., Turner II, B.L. Object-based Land-cover Classification for Metropolitan Phoenix, Arizona, using aerial photography. Int. J. Appl. Earth Obs. Geo-Inf. 2014, 33, 321-330.

Li, X., Shao, G. Object-based Land-Cover Mapping with High Resolution Aerial Photography at a County Scale in Midwestern USA. Remote Sens. 2014, 6, 11372-11390.

Li, X., Shao, G. Object-based urban vegetation mapping with high-resolution aerial photography as a single data source. Int. J. Remote Sens. 2013, 34, 771-789.

Liknes, G.C., Perry, C.H., Meneguzzo, D.M. Assessing tree cover in agricultural landscapes using high-resolution aerial imagery. J. Terr. Obs. 2010, 2, 38-55.

Lister, A.J., Scott, C.T., Rasmussen, S. Inventory methods for trees in nonforest areas in the Great Plains states. Environ. Monit. Assess. 2012, 184, 2465-2474.

Ma, L., Li, M., Ma, X., Cheng, L., Du, P., Liu, Y. A review of supervised object-based land-cover image classification. ISPRS J. Photogramm. Remote Sens. 2017, 130, 277–293.

Maxwell, A.E., Strager, M.P., Warner, T.A., Ramezan, C.A., Morgan, A.N., Pauley, C.E. Large-Area, High Spatial Resolution Land Cover Mapping Using Random Forests, GEOBIA, and NAIP Orthophotography: Findings and Recommendations. Remote Sens. 2019, 11, 1409.

Maxwell, A.E., Strager, M.P., Warner, T.A., Zégre, N.P., Yuill, C.B. Comparison of NAIP orthophotography and RapidEye satellite imagery for mapping of mining and mine reclamation. GIScience Remote Sens. 2014, 51, 301-320.

Maxwell, A.E., Warner, T.A., Vanderbilt, B.C., Ramezan, C.A. Land cover classification and feature extraction from national agriculture imagery program (NAIP) Orthoimagery: A Review. Photogramm. Eng. Remote Sens. 2017, 83, 737–747.

Meneguzzo, D.M., Liknes, G.C., Nelson, M.D. Mapping trees outside forests using high-resolution aerial imagery: a comparison of pixel-based and object-based classification approaches. Environ. Monit. Assess. 2013, 185, 6261-6275.

Meneguzzo, D.M., Lister, A.J., Sullivan, C. Summary of Findings for the Great Plains Tree and Forest Invasives Initiative. Available online: https://www.fs.fed.us/nrs/pubs/gtr/gtr_nrs177.pdf (accessed on 25 March 2020).

Montagnini, F., Nair, P.K.R. Carbon sequestration: An underexploited environmental benefit of agroforestry systems. Agroforest. Syst. 2004, 61, 281-295.

Moskal, L.M., Styers, D.M., Halabisky, M. Monitoring Urban Tree Cover Using Object-based Image Analysis and Public Domain Remotely Sensed Data. Remote Sens. 2011, 3, 2243-2262.

Myeong, S., Nowak, D.J., Hopkins, P.F., Brock, R.H. Urban cover mapping using digital, high-spatial resolution aerial imagery, Urban Ecosyst. 2003, 5, 243-256.

Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., Weng, Q. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sens. Environ. 2011, 115, 1145-1161.

Nelson, M.D., Brewer, M., Meneguzzo, D.M., Clark, K. Forests of Iowa, 2015. Available online: https://www.fs.fed.us/nrs/pubs/ru/ru_fs77.pdf (accessed on 25 March 2020).

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 2014, 148, 42-57.

O'Neil-Dunne, J.O., MacFaden, S., Royar, A. A Versatile, Production-Oriented Approach to High-Resolution Tree-Canopy Mapping in Urban and Suburban Landscapes Using GEOBIA and Data Fusion. Remote Sens. 2014, 6, 12837-12865.

O'Neil-Dunne, J., MacFaden, S., Royar, A., Reis, M., Dubayah, R., Swatantran, A. An object-based approach to statewide land cover mapping. In Proceedings of ASPRS 2014 Annual Conference, Louisville, KY, USA, 23-28 March 2014.

Pariona, A. Least Densely Populated U.S. States. Available online: worldatlas.com/articles/least-densely-populated-u-s-states.html (accessed on 7 February 2020).

Paull, D.A., Whitson, J.W., Marcotte, A.L., Liknes, G.C., Meneguzzo, D.M., Kellerman, T.A. High-resolution land cover of Kansas (2015). Fort Collins, CO: Forest Service Research Data Archive. Updated 27 November 2017. https://doi.org/10.2737/RDS-2017-0025

Platt, R.V., Schoennagel, T. An object-oriented approach to assessing changes in tree cover in the Colorado Front Range 1938-1999. For. Ecol. Manage. 2009, 258, 1342-1349.

Perry, C.H., Woodall, C.W., Liknes, G.C., Schoeneberger, M.M. Filling the gap: improving estimates of working tree resources in agricultural landscapes. Agroforest. Syst. 2009, 75, 91-101.

Qiu, X.M., Wu, S.-S., Miao, X. Incorporating road and parcel data for object-based classification of detailed urban land covers from NAIP images. GIScience Remote Sens. 2014, 51, 498-520.

Sawe, B. The 50 US States Ranked By Population. Available online: worldatlas.com/articles/us-states-by-population.html (accessed on 7 February 2020).

Schoeneberger, M.M. Agroforestry: working trees for sequestering carbon on agricultural lands. Agroforest. Syst. 2009. 75, 27-37.

St. Peter, J., Hogland, J., Anderson, N., Drake, J., Medley, P. Fine Resolution Probabilistic Land Cover Classification of Landscapes in the Southeastern United States. ISPRS Int. J. Geo-Inf. 2018, 7, 107.

USDA Forest Service, Forest Inventory and Analysis Program, Forest Inventory EVALIDator web-application Version 1.8.0.01. Available online: http://apps.fs.usda.gov/Evalidator/evalidator.jsp (accessed on 6 April 2020).

USDA National Agricultural Statistics Service, Census of Agriculture 2017. Available online: https://www.nass.usda.gov/Publications/AgCensus/2017/index.php (accessed on 6 February 2020).

USDA National Agricultural Statistics Service CropScape - Cropland Data Layer. Available online: https://nassgeodata.gmu.edu/CropScape/ (accessed on 6 February 2020).

Weather Atlas: Kansas, USA – Climate data and average monthly weather. Available online: https://www.weather-us.com/en/kansas-usa-climate (accessed 6 February 2020).

Weather Atlas: Nebraska, USA - Climate data and average monthly weather. Available online: https://www.weather-us.com/en/nebraska-usa-climate (accessed 6 February 2020).

Wickham, J., Homer, C., Vogelmann, J., McKerrow, A., Mueller, R., Herold, N., Coulston, J. The Multi-Resolution Land Characteristics (MRLC) Consortium – 20 Years of Development and Integration of USA National Land Cover Data. Remote Sens. 2014, 6, 7424-7441.

Witharana, C., Civco, D.L. Optimizing multi-resolution segmentation scale using empirical methods: Exploring the sensitivity of the supervised discrepancy measure Euclidean distance 2 (ED2). ISPRS J. Photogramm. Remote Sens. 2014, 87, 108-121.

Ye, S., Pontius, R.G., Rakshit, R. A review of accuracy assessment for object-based image analysis: From per-pixel to per-polygon approaches. ISPRS J. Photogramm. Remote Sens. 2018, 141, 137–147.

Yu, Q., Gong, P., Clinton, N., Biging, G., Kelly, M., Schirokauer, D. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. Photogramm. Eng. Remote Sens. 2006, 72, 799-811.

Zabala, A., Cea, C., Pons, X. Segmentation and thematic classification of color orthophotos over non-compressed and JPEG 2000 compressed images. Int. J. Appl. Earth Obs. Geo-Inf. 2012, 15, 92-104.

Chapter 4. Identification of windbreaks in high-resolution land cover maps from the Great Plains region of the USA using a shape-based classification model

Dacia M. Meneguzzo, Greg C. Liknes, Darci A. Paull *(draft manuscript, intended outlet: International Journal of Remote Sensing)*

4.1 Synopsis

Catastrophic dust storms carried millions of tons of topsoil away from the central United States during the 1930s. In response, United States President Franklin Roosevelt initiated a plan known as the Prairie States Forestry Project in which 217 million trees were planted on more than 30,000 farms from North Dakota to Texas, USA. The primary function of these shelterbelts, or windbreaks, is to conserve soil by redirecting wind direction and reducing wind speeds. Despite the importance of windbreaks, the current extent of the windbreak resource in the vast agricultural landscapes of the central U.S. remains largely unknown. While there have been a number of methods developed for identifying linear features in satellite or aerial imagery, few have been applied over large geographic extents. There is a growing collection of thematic land cover datasets derived from high-resolution imagery; those with a tree cover class can be exploited using shape characteristics to identify windbreaks. We present the development of a shape-based classification model that is used to identify windbreaks in a high-resolution thematic land cover dataset for Kansas, USA, which identified nearly 106,000 ha of windbreaks totaling over 50,000 km in length. An overall accuracy of 84.1% was achieved for four different windbreak classes and a non-windbreak category based on classification by an independent interpreter; overall accuracy increases to 88.8% after aggregating the windbreak classes into a single windbreak category versus a non-windbreak category.

4.2 Introduction

Prior to European settlement, the Great Plains region had an abundance of native prairie and American bison (*Bison bison*), while water and trees were scarce, especially on the uplands. Fire and the extreme climatic conditions of the plains were primarily responsible for the lack of trees; most trees occurred along streams and rivers or naturally wetter areas where they were more protected from fire (Bratton et al. 1995; Hart and Hart 1997). Since European settlement, much of the grasslands have been converted to row-crop agriculture in the eastern portion of the region while conversion to pasture/rangeland occurred in the drier western half of the region; and while fire suppression has led to an increase in trees (Hart and Hart 1997), they remain only a minor component of the landscape in terms of overall area. For example, in Kansas nearly half of the land area is now cultivated while the area of native prairie lands has been greatly reduced (Peterson et al. 2004) and forested lands occupy only about 5% of the state's total land area (USDA Forest Service, Forest Inventory and Analysis Program 2020). However, the transition to the present-day landscape was not a smooth one. The plows and farming techniques used by European settlers and overgrazing did not bode well for the fine soils of the grasslands. Overgrazing and misuse of the land coupled with severe drought and high winds in the early 1930's led to catastrophic dust storms that carried millions of tons of topsoil away from the Great Plains region of the central United States (PBS, n.d.). In response, United States President Franklin Roosevelt initiated the "Shelterbelt Project" in 1934 via an executive order that allocated money from the drought relief fund to begin planting shelterbelts in the areas most impacted by the drought ("The Establishment of a Forest Shelter Belt" 1934) and the first shelterbelt tree was planted near Mangum, Oklahoma in 1935 (USDA Forest Service 1937). Under what came to be known as the "Prairie States Forestry Project", more than 217 million trees were planted during 1935-1942 on more than 30,000 farms from North Dakota to Texas, USA and resulted in nearly 30,000 km of windbreaks (Dronen 1984). To this day, it remains "the largest

single afforestation program in U.S. history and one of the most sustained and focused efforts by the federal government to address a specific environmental challenge." (Sauer 2009, p. 2).

The primary function of these shelterbelts, or windbreaks, was to conserve soil by redirecting and reducing wind speeds (Brandle and Finch 1991). However, windbreaks provide many other ecological and economic benefits. For example, they increase crop production, reduce energy usage in farm homes, protect livestock, people, and farmsteads from harsh winds, provide wildlife and pollinator habitat, increase landscape diversity, reduce erosion, influence snow distribution, improve air quality by mitigating odors, stabilize river and streambanks, and create economic opportunities for producing firewood and other edible crops that can be sold (Brandle et al. 2004). Perhaps the function gaining most attention is that of carbon sequestration and mitigating the negative effects of climate change (e.g., Ramachandran Nair et al. 2009; Schoeneberger et al. 2012; Ballesteros-Possu et al. 2017).

Despite the importance of windbreaks, information about the location and current extent of the windbreak resource in the vast agricultural landscapes of the central U.S. remains largely unknown. The last targeted assessment of windbreaks planted during the Prairie States Forestry Project was conducted in 1954 and found that more than half (58 percent) were in poor to fair condition or had been removed (Read 1958). Other assessments in the 1970s and 1980s reported losses of windbreaks (Bratton et al. 1995). Currently, there are two national inventory programs in the United Sates that report on lands with trees and both are in the US Department of Agriculture. One is the Natural Resources Conservation Service's Natural Resource Inventory (NRI) and the other is the Forest Inventory Analysis (FIA) program in the Research and Development branch of the Forest Service. While both have collected limited information about windbreaks, these sample-based efforts use formal definitions of forest land that have size and width requirements that often exclude windbreaks from these inventories, and neither produces fine-scale geospatial products that identify and

locate windbreaks. A more recent effort, known as the Great Plains Tree and Forest Invasives Initiative (GPI) was implemented in North Dakota, South Dakota, Nebraska, and Kansas in 2008-2009 to collect data on treed lands that did not meet FIA's definition of forest land, focusing on windbreaks in particular. The GPI was a cooperative project between the USDA Forest Service and state forestry agencies in North Dakota, South Dakota, Nebraska, and Kansas designed to prepare natural resource managers in the Great Plains region for the arrival of invasive pests, mainly the emerald ash borer (*Agrilus planipennis* Fairmaire). The results of this inventory found that windbreaks comprise more than one-third of all treed lands that don't quality as forest land (Meneguzzo et al. 2018). In the second phase of GPI, referred to as GPI 2, one of the primary goals is to identify and create windbreak maps for each of these states, thus creating a geospatial inventory of the windbreak resource.

Around the same time the first GPI was carried out, a study on "working trees" by Perry et al. 2009, which included windbreaks, found that by removing the width restriction from the FIA definition of forest land, estimates of treed lands would be at least 25% higher than FIA's definitional forest land in Kansas, Nebraska (26%), South Dakota (30%), and North Dakota (38%). The results of both studies indicate that there are potentially many windbreaks in this four-state region but their extent and locations are unknown. While there may be localized efforts that keep records of new windbreak plantings, there are no ongoing coordinated region-wide efforts to inventory and monitor existing, established windbreaks. According to Ghimire et al. (2014), there is no current, statewide windbreak inventory for Kansas, nor for any other state in the U.S. to our knowledge. This research addresses this data need.

Information about windbreaks can be obtained in various ways but efforts to inventory windbreaks from the ground or map individual windbreaks by manual delineation of such features on digital aerial or satellite imagery are labor intensive and cost prohibitive (Pasher et al. 2016; Vannier and Hubert-Moy 2014). In the past, sample-based procedures were commonly used in

conjunction with remotely-sensed imagery, mainly aerial photography, to obtain summary level information (e.g., total length) about windbreaks. For example, aerial photo-based surveys of wooded strips were conducted in select areas, such as Iowa (Hartong and Moessner 1956), Kansas (Hansen 1985), and Nebraska (e.g., Schmidt and Wardle 1998). More recently, Pasher et al. (2016) applied a linear sample-based approach to estimate the total length and density of linear woody features for a large ecozone (9.5 Mha or 95,000 km2) in Canada. While sample-based methods may provide more rapid assessments of the windbreak resource as a whole, they do not create detailed maps or wall-to-wall spatial datasets that provide information about individual windbreaks, such as their location and orientation on the landscape, which informs local land management decisions.

Alternatively, automated or semi-automated remote sensing-based approaches that map windbreaks produce more detailed spatial information, e.g., length, width, and area, for individual windbreaks as well as summary information with less human effort, but require imagery with < 5 m spatial resolution. Such image-based methods often employ a geographic object-based image analysis (GEOBIA) approach, which consists of two primary processes, segmentation and classification. GEOBIA is basically a two-step process in which an image is divided into distinct homogenous segments, or image objects, that represent landscape features of interest in the first step. In the second step, the segments are assigned to a specified class based on the spectral, spatial, and relational properties of the segment. This process is said to mimic human image interpretation and is much more efficient than manual delineation of features because these processes can be automated (Wiseman et al. 2009; O'Neil-Dunne et al. 2009), and for more information on GEOBIA, see Blaschke (2010). For example, Pankiw and Piwowar (2010) and Pankiw (2013) used GEOBIA with SPOT-5 panchromatic imagery (2.5 m spatial resolution) to map shelterbelts on agricultural lands in Canada with moderate success. While no issues were noted about the adequacy of the spatial resolution, they did state that the lack of

spectral information is a limiting factor for accurate classification of shelterbelts. Similarly, Ha et al. (2019) used Sentinel imagery with a spatial resolution of 10 meters to map shelterbelts for an area in Saskatchewan and found that while their method was successful, shelterbelt mapping accuracy could be improved if imagery with a higher spatial resolution, e.g., digital aerial photos, was used.

Studies that use very high resolution (i.e., ≤ 1 meter) aerial photography or satellite imagery, in conjunction with GEOBIA have been successful in mapping windbreaks (Vannier and Hubert-Moy 2014). For example, Wiseman et al. (2009) created tree image objects from very high-resolution (62.5 cm) aerial images and then identified which objects were shelterbelts using spatial statistics. Aksoy et al. (2010) developed an automated method for detecting woody linear features from QuickBird-2 imagery with a spatial resolution of 60 cm. Woody vegetation image objects were formed and then subsequently analyzed to find those that were narrow (within a range of acceptable widths) and linear. Liknes et al. (2017) introduced a set of shape indexes specifically tailored to identify windbreaks and riparian corridors from 1-m tree cover raster thematic data. While promising, these three studies have one thing in common: the methods were applied only to geographically small study areas. Pasher et al. (2016) and Vannier and Hubert-Moy (2014) both note that semi- or fully-automated windbreak detection and mapping methods have not been implemented in an operational manner for large geographic extents. Perhaps the only exception is a study by Ghimire et al. (2014) who conducted a windbreak inventory for 14 counties in western Kansas (an area of 32,322 km^2) using GEOBIA and 1 m digital aerial photography to delineate and map windbreaks. Manual delineation of windbreaks that were missed during the classification phase was used to complete the inventory.

While semi- or fully automated linear wooded feature image detection methods are desirable, they are challenging to implement, especially over large geographic extents. For example, Piwowar et al. (2017) evaluated the use of a semi-automated GEOBIA approach to conduct a large-scale inventory of windbreaks for the agricultural region of Saskatchewan, an area of 260,000 km^2.

While suitable results could be achieved for one image, they were not repeatable when moving to different images; this problem was attributed to the complicated agricultural landscapes confusing the object-based classifier. This issue required manually adjusting the software settings for each new series of images. They found this to be more time consuming than manual delineation of windbreaks in order to produce the same results, so they opted to heads-up digitize individual windbreaks on high-resolution digital aerial photos instead. A windbreak inventory was completed but they note that this approach was intensive in terms of labor, time, and budget.

The above studies have applied GEOBIA directly to high resolution imagery and have noted the challenges of doing so, an alternative option is to identify windbreaks in readily available thematic maps of tree cover, in particular those derived from high resolution imagery. As high-resolution land cover datasets for large extents, whether national (e.g., Robinson et al. 2019) or regional (e.g., O'Neil-Dunne et al. 2014; Basu et al. 2015; St. Peter et al. 2018; Maxwell et al. 2019; Chesapeake Conservancy 2016, Paull et al. 2017; Kellerman et al. 2019) become more available, methods to extract windbreaks are needed. One property of windbreaks that makes them easily recognizable is their distinct linear shape. Shape is one of the fundamental elements of human photo interpretation and features with very distinctive shapes can be identified using this information alone (Olson 1960). For example, van der Werff and van der Meer (2008) used geometric indexes to classify various types of water bodies into separate categories. This is a good example of using shape information to distinguish among objects that are the same land cover type or in the same thematic class, which is one of our objectives.

Shape metrics are advantageous for classification because, unlike spectral attributes, they are not impacted by color or light conditions (Benz et al. 2004). However, a lingering challenge in windbreak mapping is identifying multiple-leg windbreaks. While windbreaks frequently occur as a single linear leg oriented in a north-south or east-west direction, those with more than one leg are

also common. Such configurations include two legs that occur at right angles to each other shaped like the capital letter “L,” or those with three or more legs that can occur in various shapes, such as forming what looks like a “U” shape or hollow square from an aerial view. Both Liknes et al. (2017) and Ha et al. (2019) note difficulty with mapping multiple-leg windbreaks that are in the shape of an “L”. The objective of this study was to develop a practical, operational method for identifying both single and multiple-leg windbreaks from high-resolution (1m) land cover maps that would be scalable for large regions. Utilizing shape indexes developed by Liknes et al. (2017) and a suite of shape statistics derived from eCognition software, a windbreak classification model was developed using training samples from a six-county area in Kansas, USA. The model was then applied to high-resolution thematic tree cover data for the entire state of Kansas (Paull et al. 2017), and total number, area and length of windbreaks was calculated for all counties. We note that various definitions of “windbreak” exist (e.g., Brandle et al. 2004, USDA National Agroforestry Center, 2012), but for this study we use the term in reference to landscape features comprised of multiple trees of either natural or planted origin, and arranged in one or more connected legs that are of a regular linear shape, for which the likely intended land use is to reduce negative effects of wind on adjacent land uses.

4.3 Methods

4.3.1 Study Area

The state of Kansas, USA, lies in the center of the United States and is divided into 105 counties (Figure 4.1). Currently, Kansas covers an area of nearly 213,100 km^2 and is dominated by land uses related to agriculture. Historically, the landscape was much different. The native vegetation of Kansas was made up various prairie grasses with patches of woodlands; forests moved in when there was an ample supply of moisture and retreated with drought and wildfire (Bratton et al. 1995). The topography gradually rises from east to west and the climate is one of extremes with recurring droughts. With European settlement came dams,

wildfire suppression and mass conversion of grasslands to cropland and intensive grazing by cattle (Bratton et al. 1995). The combination of severe droughts and soil mismanagement let to the catastrophic dust storms mentioned above (Sauer 2009). The landscape today is one of intensively managed monocultures (Bratton et al. 1995).

Currently, the top three land use classes in Kansas are grassland/pasture (46%), corn (19%), and soybeans (10%) (USDA National Agricultural Statistics Service Cropscape Cropland Data Layer) and only a small percentage (5 percent) of the state is classified as forest land (USDA Forest Service, Forest Inventory and Analysis Program). Kansas is one of the least densely populated states in the U.S. with a population density of about 14 people per square kilometer (Pariona 2017).

The model development study area is comprised of six counties in Kansas, USA, including Chase, Gove, Jackson, Seward, Stafford, and Wilson, that make up a cumulative area 11,702 km^2 in size, or about 5.5% of the state's total area (Figure 4.1). These particular counties were selected because they provide a representative sample of the range of tree cover patterns and densities across Kansas (Figure 4.2[a-f]). Tree cover as a percent of total area by county ranges from a low of 0.3 percent in Seward County (Figure 4.2a) to a high of 19.8 percent in Wilson County (Figure 4.2f). Agricultural land uses such as grassland/pasture and cropland make up at least 79.2 percent of all land area in each of the counties (Figure 4.3) (USDA National Agricultural Statistics Service).

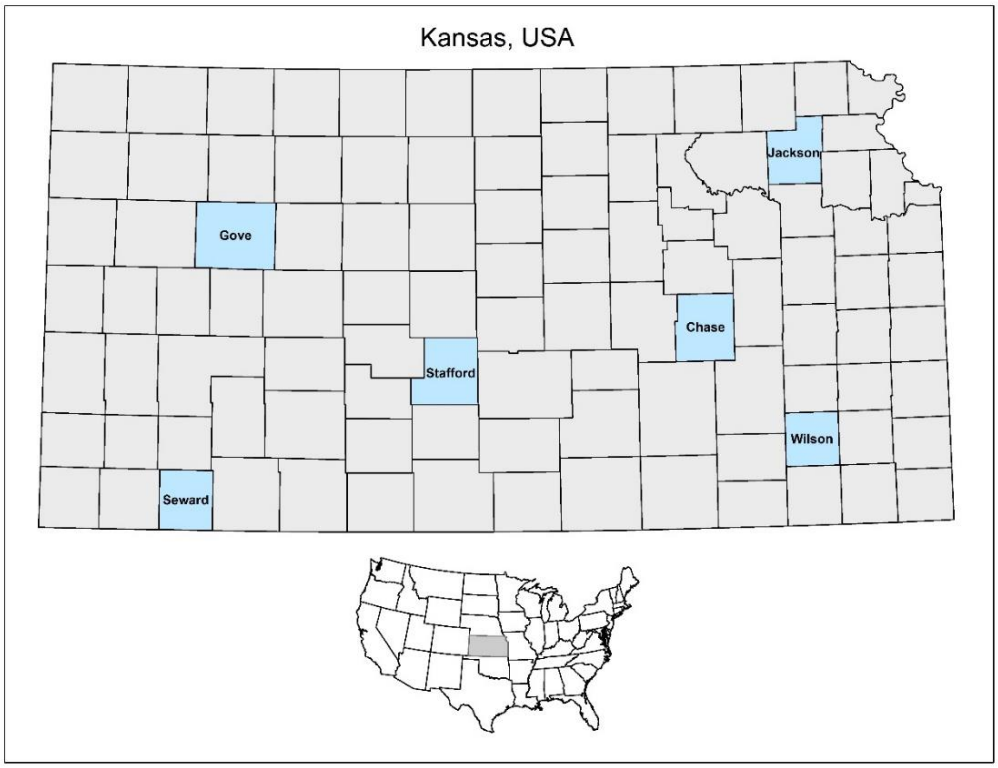
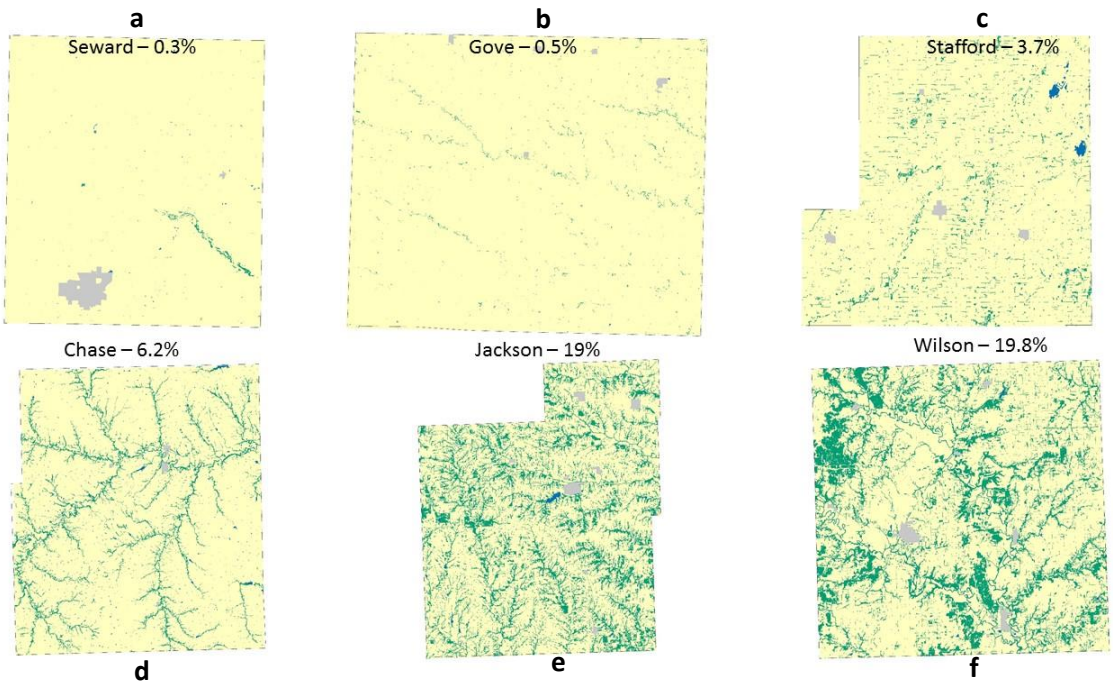


Figure 4.1. Six-county windbreak classification model development study area in Kansas, USA.



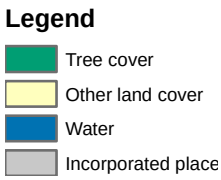


Figure 4.2. Windbreak classification model development study area consisting of six counties **(a-f)** that represent a range of tree cover densities (labeled at the top of each county) and patterns in Kansas, USA.

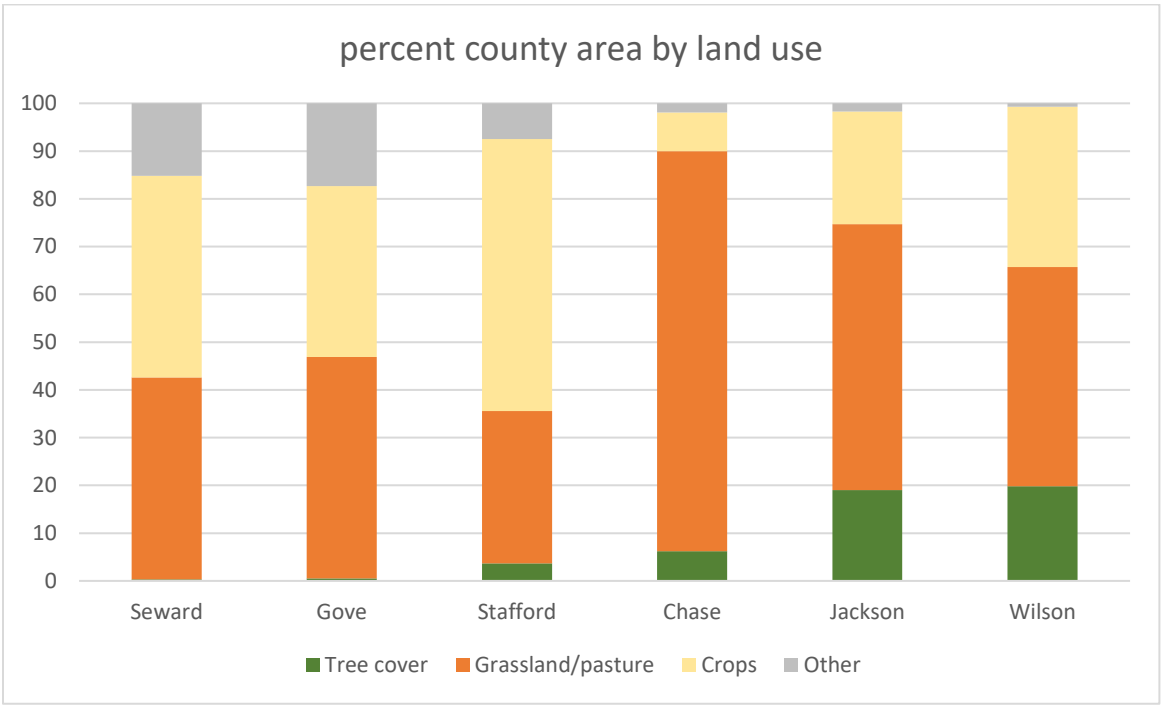


Figure 4.3. Percentage of county area by general land use for six counties in Kansas, USA. (Data source: USDA National Agricultural Statistics Service CropScape - Cropland Data Layer. Available online: https://nassgeodata.gmu.edu/CropScape/)

4.3.2 High-resolution Tree Cover Map

Tree cover data for Kansas were obtained from the statewide 1 m land cover dataset entitled *High-resolution land cover of Kansas (2015)* (Paull et al., 2017), which was produced following the methods of Liknes et al. (2010) and

Meneguzzo et al. (2013). This dataset has an overall accuracy of 97.7% and the user's accuracy of the tree cover class is 98.6% (see Table 3.3 in Chapter 3, section 3.4.2.1). These land cover data were derived from 1-meter aerial imagery from the National Agriculture Imagery Program (NAIP) and consist of four land cover classes (tree cover, other land cover, water, or city/town); the dataset is distributed as 105 county-level thematic raster datasets. For the six-county modeling area, the four-class county-level land cover datasets were converted to binary raster files where tree cover pixels were set to foreground pixels (value = 1) and all other pixels were given the background value of 0. A script was written in Python to batch convert the remaining 99 county-level land cover datasets to the binary (tree/not tree) raster files so the model could later be applied to all counties in the state. We also note that incorporated areas were excluded from the land cover dataset so tree cover in these areas were not included in this study; the focus is rural lands only.

4.3.3 Geographic Object-based Image Analysis (GEOBIA) Approach

Geographic Object-based image analysis (GEOBIA) is essentially a two-step process: image segmentation and classification. Segmentation divides the input image into segments, also called image objects, which become the classification units (Blaschke 2010). While this approach is frequently used to create various types of land use/land cover products from digital aerial photography or satellite images, we applied it in a less traditional fashion to create segments from the foreground (tree cover) pixels in the binary raster data described above. The tree cover segments were then classified into various windbreak categories.

4.3.3.1 Segmentation

The county-level binary raster files for the six counties in the model development study area were segmented using eCognition® Developer software version 9. An iterative process was used to determine which segmentation

parameters that produced segments that delineated the patches of tree cover by following the boundaries of the tree cover pixels in the binary raster maps. In particular, we used the chessboard segmentation algorithm with object size = 10 followed by the spectral difference algorithm with the maximum spectral difference = 0.9. This process created segments that represented each distinct group of tree pixels in the raster data. The resulting tree segments and selected image object attributes were exported in vector (i.e., shapefile) format. The same segmentation routine and parameters were applied to the remaining 99 counties in the state.

4.3.3.2 Feature Selection

Liknes et al. (2017) described three shape indexes (see Table 4.1), with each capturing a different aspect of common windbreak shapes. The straight-and-narrow feature index (SNFI) is the difference in the count of single-width pixel kernels oriented east/west and north/south that fit within a particular map object and normalized by the sum of counts for both kernels. We utilized SNFI in this study with both 37-m and 74-m kernel lengths which approximate common single-leg windbreak widths observed in the Great Plains region. In addition, we included a non-normalized version of SNFI (e.g., “Foc37dif” in Table 1) for both kernel lengths and also the mean pixel counts for east/west (e.g., “FocX37”) and north/south (e.g., “FocY37”) separately as attributes.

The second index is windbreak sinuosity (“Sinuous”), which is the ratio of the length traveled along half of a map image object’s perimeter to the diagonal length of the object’s bounding box. It quantifies the deviation of the image object from rectangular; we’ve observed that more curvilinear features are more commonly associated with sinuous-shaped riparian forest buffers found along winding streams in the Great Plains region. Lastly, the area index (“Area_Ind”) is the proportion of a map object’s bounding box filled by the map object. This index was designed with the goal of identifying L-shaped windbreaks because they should have smaller values relative to straight-line windbreaks that occupy most

of the bounding box area. We used both windbreak sinuosity and area index as described by Liknes et al. (2017). The three shape index values were produced for each tree image object within the eCognition Developer software using various customized functions. This is an improvement from our previous research (Liknes et al., 2017) that required additional outside processing using ArcGIS® software, thus making the current process much more efficient.

In addition, there are a number of shape metrics that are readily available in the eCognition Developer software; 15 were selected and generated for all tree segments within each county and exported as attributes in the shapefiles described in section 4.3.3.1 (Table 4.1). Some of the notable shape properties include asymmetry, shape index, and density. Wiseman et al. (2009) found these metrics to be particularly useful for distinguishing single-leg windbreak image objects from other tree cover segments for a small study area in Manitoba, Canada. We note there conceptually is some overlap with the indexes from Liknes et al. (2017); for example, the area index described above is very similar to the “Rectangular Fit” metric available in eCognition.

Table 4.1. Shape indexes and geometric image object attributes used as predictor variables in developing a windbreak classification model for Kansas (and Nebraska), USA. Shape-based information are for image objects representing tree cover only.

Attribute	Short name
Shape Index (Liknes et al. 2017)	
Focal Difference 37 (37-m width)	Foc37dif
Focal Difference 74 (74-m width)	Foc74dif
Focal X direction (37-m width)	FocX37
Focal Y direction (37-m width)	FocY37
Focal X direction (74-m width)	FocX74
Focal Y direction (74-m width)	FocY74
Straight and Narrow Feature Index (37-m width)	SNFI37
Straight and Narrow Feature Index (74-m width)	SNFI74
Windbreak Sinuosity Index	Sinuous
Area Ratio Index	Area_Ind
Shape Metric from eCognition	
Area	Area_Pxl
Asymmetry	Asym
Border Index	Border
Compactness	Compact
Density	Density
Elliptic Fit	Elliptic
Length of Main Line	Length
Length/Width	LenWid
Main Direction	Main_dir
Radius of Largest Enclosed Ellipse	RadLarge
Radius of Smallest Enclosing Ellipse	RadSmall
Rectangular Fit	RectFit
Roundness	Round
Shape Index	ShapeInd
Width	Width

4.3.3.3 Supervised Classification and Random Forests™ Classification Model Development

We used supervised classification with a good quality, representative sample of windbreaks that were collected by an interpreter from the shapefiles containing the tree segments and their associated attributes; NAIP imagery from 2015 served as the background. Samples were collected in spatially balanced manner throughout the county and only delineated tree cover polygons from the input map were considered, i.e., non-delineated tree cover (including windbreaks) was not included or manually delineated by the interpreter. While windbreaks may seem like a relatively easy feature to identify, they occur in a variety of configurations and orientations to serve specialized ecosystem functions (Figure.4.4). High-quality, representative training samples were collected for the five windbreak categories (Table 4.2) from each of the six counties in the model development study area. The goal was to collect at least 15 samples of each windbreak category from each county for a total of at least 50 samples per category for the model development study area. However, it was not always possible to collect a minimum of 15 samples of each category in each county, especially in counties with a very low percentage of tree cover. If there weren't enough high-quality samples available to meet this goal, as many good quality samples as possible were collected and additional samples could be collected from other counties if necessary. After some initial investigation of the classification output, it was decided that the total minimum number of samples of each category be increased to 100.

Table 4.2. Description of the five windbreak categories, corresponding figure number, and number of training data samples collected to develop a windbreak classification model from a six-county study area in Kansas, USA.

Windbreak Category Code	Shape/configuration	Corresponding letter in Figure 4	Number of samples collected
1	North-south windbreak	a	102
2	East-west windbreak	b	102
3	L-shape windbreak	c	103
4	Complex windbreak (more legs/complex shape than categories 1-3)	d	103
5	Other/non-windbreak	e	190

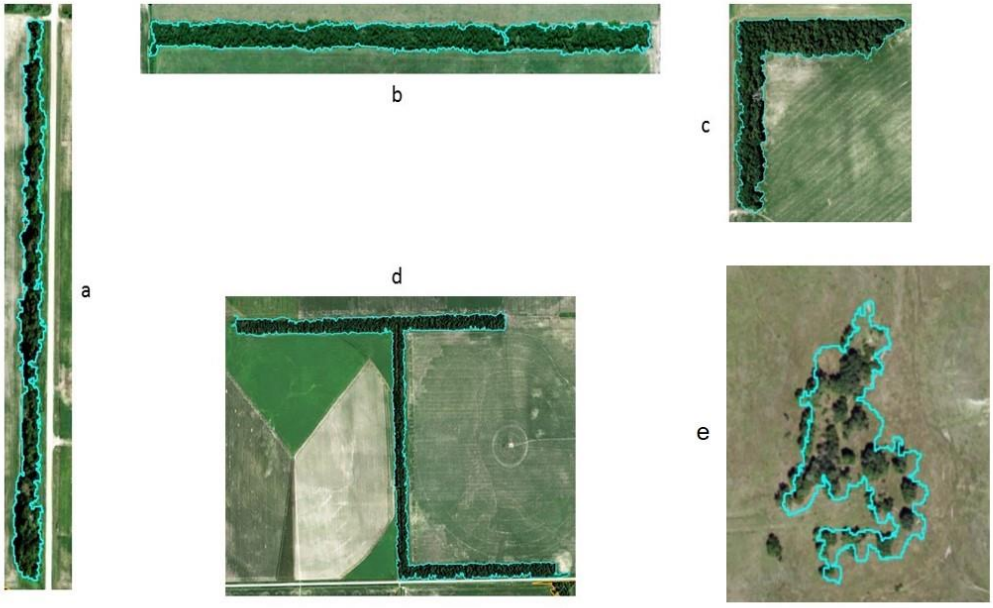


Figure 4.4. Examples of the four windbreak categories (a-d) and a non-windbreak category representing non-windbreak tree features (e) collected as training data for windbreak classification model development in Kansas, USA.

The training data samples from each county were compiled into one training data file. These data were then used to build a shape-based windbreak classification model using the Random Forests (RF™) algorithm in the open source R statistical computing environment. Tree image objects for the six counties in the model development study area were labelled by the model as one of the five windbreak categories: north-south, east-west, L-shape, complex, or other (non-windbreak) based on their shape-related attributes.

4.3.4 Linear Discriminant Analysis and Variable Importance

Basic exploration of the training data and their associated shape metrics was conducted in the R v.3.3.1 statistical computing environment. A multi-variate tool, linear discriminant analysis (LDA), was applied to the training data. LDA is used to maximize the separability between the input categories (Martínez and Kak 2001). The output LDA plot allows us to visualize and assess how well the five windbreak categories can be separated using the shape metrics as predictor variables. We also used the “mean decrease Gini” variable importance measures provided by Random Forests to identify which predictor variables are the most important in predicting the response variable given the large number of predictor variables. Predictor variables are ranked in order of importance and displayed using a variable importance plot.

4.3.5 Accuracy Assessment

4.3.5.1 Out-of-bag Error Assessment

One of the advantages of the RF algorithm is that it provides an unbiased estimate of error based on randomly selected observations that are withheld during model development, referred to as the out-of-bag (OOB) sample (Breiman and Cutler, n.d.). The OOB observations can be used to evaluate classification accuracy (Breiman 2001a; Breiman and Cutler, n.d.; Gislason et al., 2006; Cutler et al., 2007) because, in essence, they are cross-validation data that provide an

estimate of accuracy (Cutler et al., 2007). This information was used to assess accuracy for the six-county study area by calculating class-level producer's and user's accuracies as well as overall accuracy from an error matrix.

4.3.5.2 Independent Accuracy Assessment

In addition to the OOB assessment, we also conducted an independent accuracy assessment using high-resolution aerial photography as the reference data since there are no existing high-resolution windbreak products available. Because the windbreak map was created from a land cover dataset derived from publicly available aerial imagery acquired in 2015 from the National Agriculture Imagery Program, the same aerial imagery is used as the reference data.

A simple random sampling design likely would have under-sampled uncommon classes, so a stratified random sampling design where a minimum number of samples are selected for each category (Lunetta 2004) was implemented. A minimum of 50 samples for each category is a general rule for assessing accuracy using an error matrix (Lunetta 2004; Congalton and Green 2009). It was determined that 50 samples from categories 1-4 would be randomly selected from each county and assigned a class label by an interpreter. Since the non-windbreak category had many more samples compared to the other categories, 100 samples from this category were randomly selected for each county (Lunetta 2004). The labels assigned by the interpreter were compared to those assigned by the classification model to assess agreement between the aerial photo reference data and the windbreak classification model results. Because some of the counties had very little tree cover, there were fewer than 50 classified image objects available in some windbreak categories. In Seward County, for example, only seven of the image objects in the county were classified as complex windbreaks (class 4). In those cases, all of the samples were included. All of the samples from each county were pooled together for a total of 1,666 samples in the photo-interpreted accuracy assessment. Similar to the OOB assessment, this information was used to assess accuracy for the six-county study area by

calculating class-level producer's and user's accuracies as well as overall accuracy from an error matrix.

4.3.6 Windbreak Area and Length Estimates

While area is a commonly used metric for measuring land cover or use, total length is also a desired attribute when quantifying narrow linear features like windbreaks. Length is a logical metric for inventory and monitoring of the linear resource and it is a necessary attribute for quantifying the ecosystem services provided by windbreaks. Thus, a distance measure like length is more appropriate to describe linear features and can be used to monitor changes over time. For example, Amichev et al. (2015) notes not only the lack of shelterbelt maps but also data describing the total length of shelterbelts in the agricultural areas of Saskatchewan. Currently, a similar lack of data exists in the U.S. In the past, intermittent sample-based inventories produced estimates of total area and length of windbreaks for the area under consideration. For example, a study by Hansen (1985) estimated the area and total length of wooded strips for the state of Kansas. Length estimates for the windbreak categories and area estimates for windbreaks as well as the non-windbreak category were provided by the length and area geometric attributes that were exported from eCognition in vector format (refer to Table 4.1). Obtaining the length of the L-shape windbreaks required summing the length and width attributes. Determining the length of the complex windbreaks proved to be more challenging because some of these windbreak systems have three or more legs and/or very unique shapes. After exploring various metrics and options in eCognition and ArcGIS, we found that using the centerline vectorization method with the ArcScan extension in ArcGIS was a successful method for creating a length metric for these more complex features. Ideally, the area and length estimates would be produced from reference data (Olofsson et al., 2014), but no comparable windbreak datasets exist for Kansas (Ghimire 2014) that could be used as reference data, nor did we have the time or resources to create one.

4.4 Results

4.4.1 Windbreak Classification Model Development

A windbreak classification model was developed using a collective set of 508 training data samples and applied to the six county study area. These preliminary results were assessed by the OOB error report, LDA plot, and visual examination of the classification output. All three methods revealed frequent confusion amongst L-shape windbreaks, complex windbreaks, and non-windbreaks, categories 3, 4, and 5, respectively. To remedy this, the training data were refined by removing samples that were not accurate depictions of the distinct windbreak shape categories and additional samples of the “other” category were collected because of the high variability of shapes in this category. The new training data set had a total of 600 samples (refer to Table 4.2). A new windbreak classification model was built and applied to the study area. The same three assessment methods were used and the new results showed improvement, especially in categories 4 and 5. The new model was applied on a county-by-county basis to the remaining 99 counties in Kansas. The Mean Decrease in Gini variable importance measures were also examined via a variable importance plot to see which predictor variables are most important to the model.

4.4.2 Linear Discriminant Analysis and Variable Importance

LDA was applied to the collective set of 600 training data samples. We did get a collinearity warning so there is correlation among some of the input variables; therefore the data fail to meet the assumptions of the LDA model. However, we show the resulting plot here to provide a visual assessment of potential separability between the windbreak categories. The LDA plot shows that several of the SNFI-related attributes (Foc37dif, FocX74, and Foc74dif) as well as density and rectangular fit are important variables for separating the data (Figure 4.5). In particular, there is good separation between windbreak categories 1 and

2 (north-south and east-west oriented windbreaks, respectively); they separate very well from each other and from the other three categories. The proportion of trace (“LD1” in Figure 4.5) indicates that we obtained about 58% separation amongst the categories by the first (best) discriminant function.

There is a lot of overlap among categories 3 and 4 in particular, and, to a lesser extent, among the non-windbreak category and the multiple-leg categories. This, along with the collinearity warning, indicates that we need a more powerful, nonparametric classifier such as Random Forests, to handle the correlated data (Belgiu and Drăguţ 2016) as well as the nonlinear relationships among the input predictor variables (Yan and Roy 2015). We note here that LDA was not used any further in this study; the classification model was built using RF.

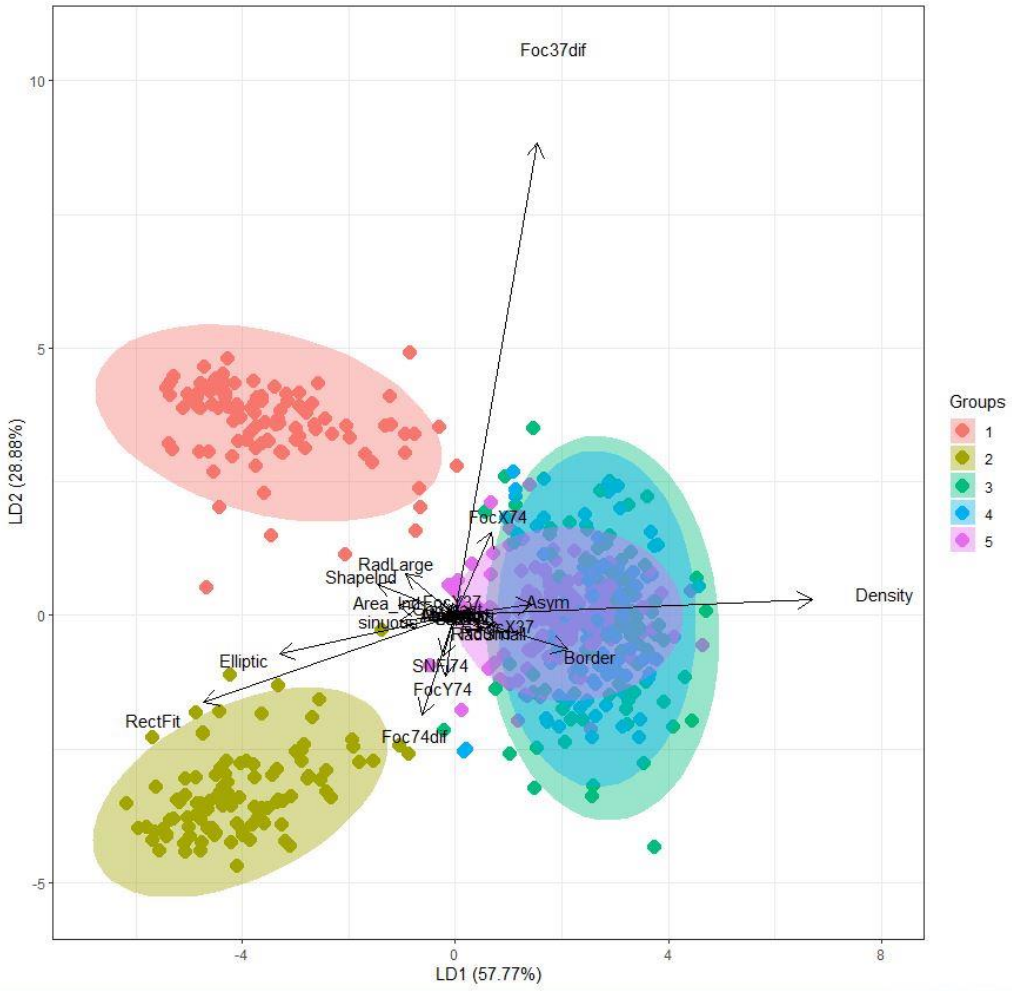


Figure 4.5. Linear discriminant analysis plot showing the separability of windbreak categories (1-4) and a non-windbreak category (5) using a variety of shape-based predictor variables.

The variable importance plot produced by Random Forests based on the training data (Figure 4.6) indicates that the top ten most important predictor variables are: Straight and Narrow Feature Index (74-m width), length to width ratio, focal difference (37-m width), windbreak sinuosity index, asymmetry, Straight and Narrow Feature Index (37-m width), roundness, radius of smallest enclosing ellipse, focal difference (74-m width), and density. There is overlap with the LDA findings, namely the focal difference (37-m and 74-m widths) and density predictor variables.

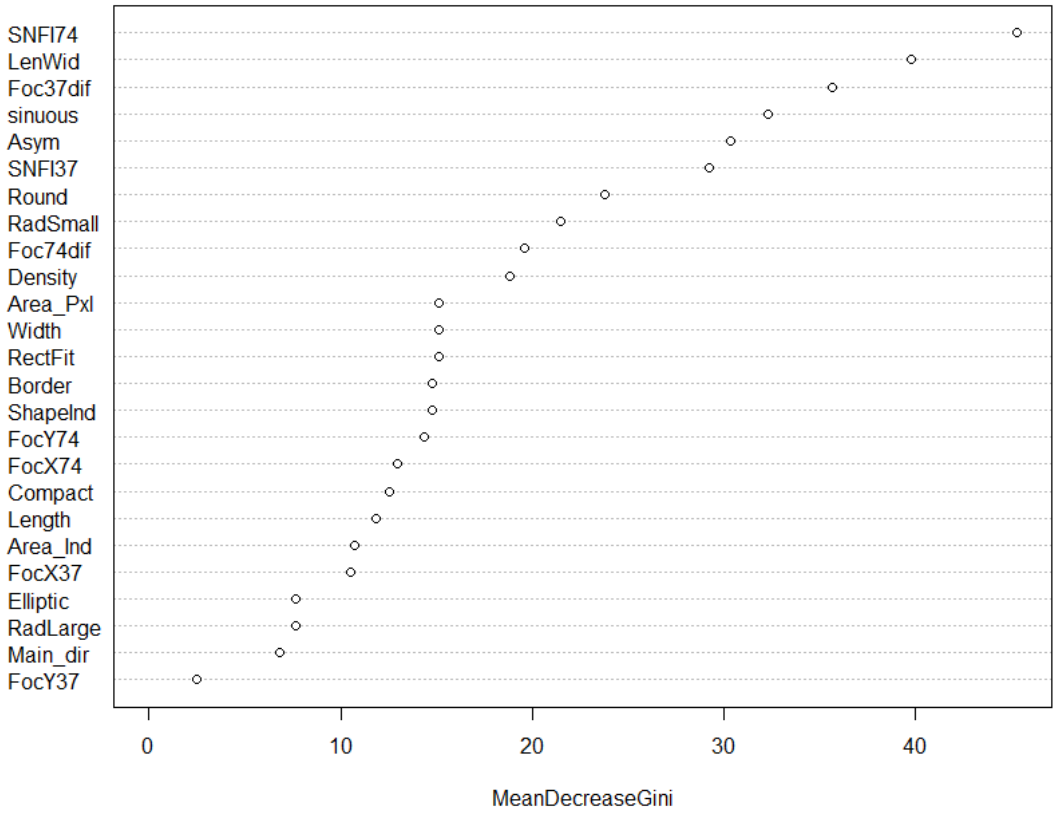


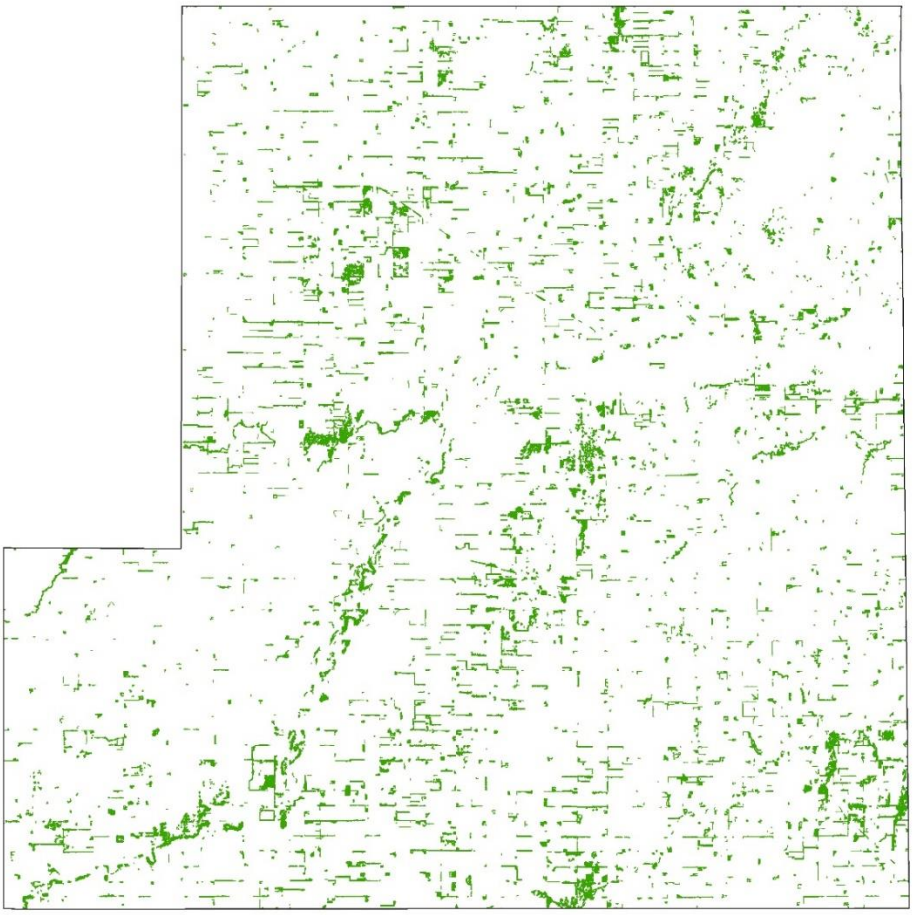
Figure 4.6. Variable importance plot ranking shape-based predictor variables in order of importance (higher mean decrease Gini values indicate higher importance) for a windbreak classification model developed for Kansas, USA.

4.4.3 Windbreak Classification Model Application

A statewide windbreak dataset was produced using the windbreak classification model developed from the six county study area. A map of tree cover in Stafford County before and after applying the windbreak classification model are shown in Figure 4.7. There are more than 28,000 tree image objects in the six-county study area and more than one-half million statewide. In the study area, 20 percent of the tree cover image objects were classified as one of the windbreak categories: 7 percent were classified as north-south windbreaks, 7 percent as east-west windbreaks, approximately 4 percent as L-shaped windbreaks, and 2 percent were classified as complex windbreaks; statewide percentages are

similar (Table 4.3). North-south windbreaks were nearly equivalent to east-west windbreaks in the study area but occurred more frequently throughout the state. Complex windbreaks were the least common windbreak category in all counties of the study area as well as at the state level.

(a)



(b)

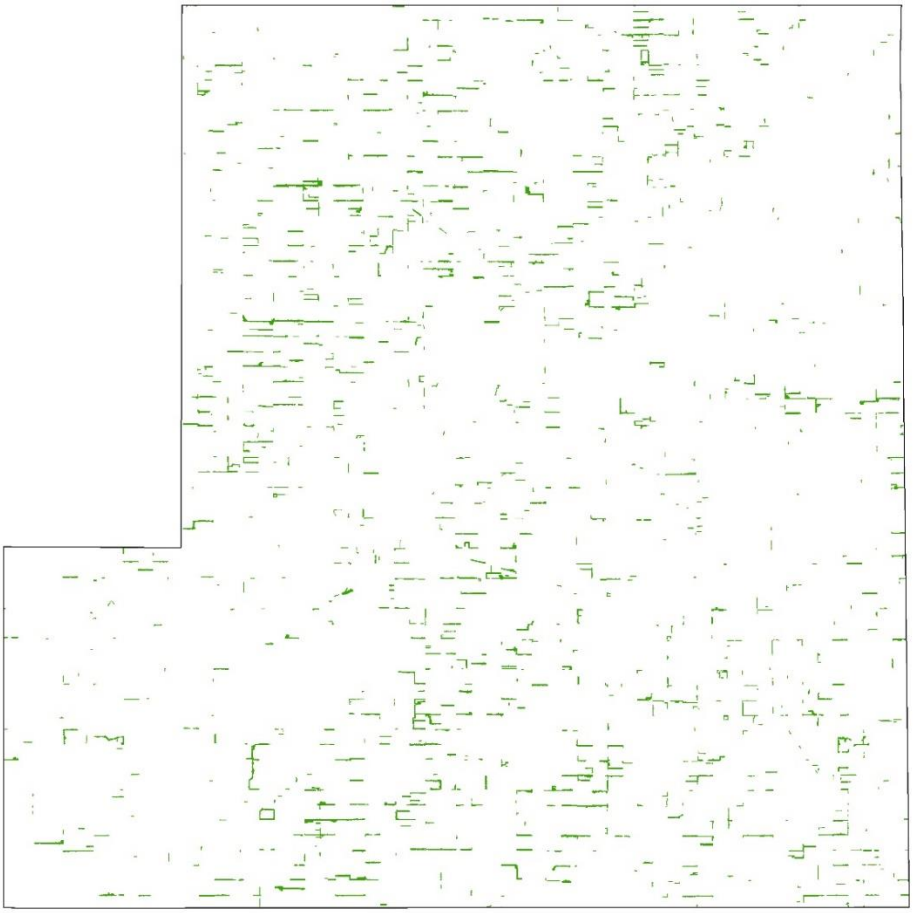


Figure 4.7. Tree cover map with 1-m resolution for Stafford County, Kansas, USA (a) and resulting windbreak map for Stafford County after applying a windbreak classification model based on shape attributes developed for Kansas, USA (b).

Table 4.3. Number of tree cover image objects classified as one of four windbreak categories or a non-windbreak category for a six county study area and the state of Kansas, USA, based on a windbreak classification map derived from a 1 m thematic raster map of tree cover. Percentage of the total number of image objects shown in parentheses.

County	Total number of tree cover image objects	Number of image objects classified as Class 1 (north-south windbreak)	Number of image objects classified as Class 2 (east-west windbreak)	Number of image objects classified as Class 3 (L-shape windbreak)	Number of image objects classified as Class 4 (complex windbreak)	Number of image objects classified as Class 5 (other)
Chase	4,358	268 (6.1%)	194 (4.5%)	168 (3.9%)	101 (2.3%)	3,627 (83.2%)
Gove	2,327	80 (3.5%)	198 (8.5%)	112 (4.9%)	20 (0.9%)	1,917 (82.3%)
Jackson	9,323	674 (7.3%)	357 (3.8%)	300 (3.3%)	161 (1.7%)	7,831 (83.9%)
Seward	506	23 (4.7%)	37 (7.1%)	28 (5.7%)	7 (1.4%)	411 (81.0%)
Stafford	5,197	280 (5.3%)	795 (15.4%)	169 (3.3%)	100 (1.9%)	3,853 (74.1%)
Wilson	6,784	679 (10.2%)	421 (6.2%)	334 (5.0%)	206 (3.0%)	5,144 (75.6%)
Study Area Total	**28,495**	**2,004 (7.0%)**	**2,002 (7.0%)**	**1,111 (3.9%)**	**595 (2.1%)**	**22,783 (79.8%)**
State Total	**542,589**	**41,656 (7.7%)**	**35,588 (6.6%)**	**27,195 (5.0%)**	**13,576 (2.5%)**	**424,574 (78.2%)**
County-level range (low – high)	**192 – 13,620**	**14 – 1,277**	**15 – 1,556**	**12 - 972**	**0 – 587**	**108 - 10,886**

Morton County, located in the southwest corner of the state, had the fewest classified windbreaks with a total of 46. Wyandotte County only had a total of five but isn't considered because most of the county is urban class, which was excluded from windbreak mapping. Reno County, located in south-central Kansas, had the most classified windbreaks with 4,150 and is the only county to exceed 4,000 windbreaks. McPherson County had the highest number of north-south windbreaks with 1,277 while Reno County had the most east-west (1,556)

and complex windbreaks (587). The highest occurrence of L-shape windbreaks occurred in Dickinson County with 972.

4.4.4 Assessing Map Accuracy

4.4.4.1 Out-of-bag Error Assessment

The OOB sample for the six county study area was one of the methods used to assess accuracy (Table 4.4). The north-south and east-west windbreak classes had very high accuracy while the accuracies for the L-shape and complex categories were lower. The results of the OOB accuracy assessment have a pattern similar to that shown in the LDA plot in terms of confusion, or lack of separability, amongst categories 3, 4, and 5. The user's and producer's accuracies for the single-leg windbreak categories are 98.1% or higher while the multiple-leg categories (L-shape and complex) range from a low of 73.3% to a high of 79.6%. The non-windbreak category has a user's accuracy of 88.9% and a producer's accuracy of 90.9%. Overall accuracy is 88.5% (Table 4.4).

Table 4.4. Accuracy assessment of the windbreak classification results for a six-county study area in Kansas, USA, based on the out-of-bag sample containing 600 observations.

	Reference						
Classification	**North-south wind-break**	**East-west wind-break**	**L-shape wind-break**	**Complex wind-break**	**Other (non-windbreak)**	**Row Total**	**User's Accuracy**
North-south windbreak	101	0	0	0	1	102	99.0%
East-west windbreak	0	102	0	0	0	102	100.0%
L-shape windbreak	0	0	82	14	7	103	79.6%
Complex windbreak	0	1	16	77	9	103	74.8%
Other (non-windbreak)	0	1	6	14	169	190	88.9%
Column Total	101	104	104	105	186	600	
Producer's Accuracy	100.0 %	98.1%	78.8%	73.3%	90.9%		
Overall Accuracy	**88.5%**						

4.4.4.2 Independent Accuracy Assessment

The results for the photo-interpreted windbreak classification accuracy assessment are presented in the following error matrix (Table 4.5). The user's accuracies are lower than those in the OOB error assessment but follow a similar pattern for the four windbreak categories, while the results for the non-windbreak category are both about 90%. The producer's accuracies are lower with the exception of the L-shape windbreak category and the complex category has about the same results in both error matrices. The user and producer accuracies

for the two single-leg windbreak categories were 88.0% or higher while the more complex windbreaks have lower accuracies, around 70%, with the exception of the L-shape producer's accuracy mentioned above. Overall accuracy is 84.1%

Table 4.5. Accuracy assessment of the photo-interpreted windbreak classification results for a six-county study area in Kansas, USA.

	Reference						
Classification	**North-south wind-break**	**East-west wind-break**	**L-shape wind-break**	**Complex wind-break**	**Other (non-windbreak)**	**Row Total**	**User's Accuracy**
North-south windbreak	250	0	2	0	22	274	91.2%
East-west windbreak	0	264	5	4	13	286	92.3%
L-shape windbreak	11	6	196	35	31	279	70.3%
Complex windbreak	2	5	21	155	44	227	68.3%
Other (non-windbreak)	21	11	14	18	536	600	89.3%
Column Total	284	286	238	212	646	1,666	
Producer's Accuracy	88.0%	92.3%	82.4%	73.1%	83.0%		
Overall Accuracy	**84.1%**						

Most of the confusion for the single-leg windbreak categories occurs with the non-windbreak category versus with the L-shape and complex categories, and the main source of confusion with the non-windbreak category is north-south windbreaks. Examination of the confused samples indicated that riparian forest buffers often have a shape that closely resembles windbreaks with a north-south orientation, so, for example, a north-south windbreak that deviates from linear

may be incorrectly classified as a non-windbreak and vice versa: a relatively linear riparian forest buffer may be labelled as a north-south windbreak. Photo interpretation or ancillary data such as hydrography dataset of stream lines could be used to assign these groups of trees to the non-windbreak category, or they could be assigned to a new “riparian buffer” class for future assessments of other tree functional types.

Overall, most confusion occurred between the L-shape and complex windbreak categories; there was also a fair amount of confusion between these more complex windbreak classes and the non-windbreak category. By combining these into one multiple-leg category, the user’s accuracy for the new class is 80.4% and the overall accuracy increases to 87.5% (Table 4.6). If we take things a step further and compile all of the windbreak categories into one class so we have only two classes: windbreak and non-windbreak, overall accuracy increases to 88.8%. This may be sufficient if the user is only interested in windbreaks versus non-windbreak groups of trees and not the orientation or specific shape of the windbreaks.

Table 4.6. Results of the photo-interpreted windbreak classification results for a six-county study area in Kansas, USA, in which the L-shape and complex windbreak categories have been combined into one multiple-leg windbreak category.

	Reference					
Classification	**North-south wind-break**	**East-west windbreak**	**Multiple-leg windbreak**	**Other (non-windbreak)**	**Row Total**	**User's Accuracy**
North-south windbreak	250	0	2	22	274	91.2%
East-west windbreak	0	264	9	13	286	92.3%
Multiple-leg windbreak	13	11	407	75	506	80.4%
Other (non-windbreak)	21	11	32	536	600	89.3%
Column Total	284	286	450	646	1,666	
Producer's Accuracy	88.0%	92.3%	90.4%	83.0%		
Overall Accuracy	**87.5%**					

4.4.5 Windbreak Area and Length Estimation

Area estimates for the windbreak (and non-windbreak) categories and length of the windbreak categories were obtained from county-level classification output datasets. Map-based area estimates from the high-resolution windbreak datasets indicate that there are more than 1.5 million ha of tree cover in Kansas, of which 105,840 ha, or about 7%, are classified as windbreaks (Table 4.7). Haskell County had the smallest area of tree cover classified as windbreaks with

17.7 ha while Reno County had the largest area with 4,507.9 ha. The average area per county is just over 1,000 ha.

Table 4.7. Area (ha) estimates for four windbreak categories and a non-windbreak category derived from a windbreak classification map for a six county study area and the state of Kansas, USA. Percentage of the total area shown in parentheses.

County	Total area of tree cover image objects (ha)	Area of image objects classified as north-south windbreak (ha)	Area of image objects classified as east-west windbreak (ha)	Area of image objects classified as L-shape windbreak (ha)	Area of image objects classified as complex windbreak (ha)	Area of image objects classified as other (ha)
Chase	12,356.0	133.8 (1.1%)	124.4 (1.0%)	90.5 (0.7%)	405.1 (3.3%)	11,602.2 (93.9%)
Gove	1,446.5	36.4 (2.5%)	76.2 (5.3%)	54.0 (3.7%)	33.3 (2.3%)	1,246.6 (86.2%)
Jackson	32,209.2	401.4 (1.2%)	142.9 (0.4%)	179.3 (0.6%)	582.0 (1.8%)	30,903.6 (95.9%)
Seward	523.0	6.4 (1.2%)	9.7 (1.9%)	10.9 (2.1%)	3.9 (0.7%)	492.1 (94.1%)
Stafford	7,562.4	129.2 (1.7%)	1,144.3 (15.1%)	255.5 (3.4%)	517.3 (6.8%)	5,516.1 (72.9%)
Wilson	29,481.6	345.5 (1.2%)	174.3 (0.6%)	214.5 (0.7%)	641.9 (2.2%)	28,105.4 (95.3%)
Study Area Total	**83,578.7**	**1,052.7 (1.3%)**	**1,671.9 (2.0%)**	**804.7 (1.0%)**	**2,183.4 (2.6%)**	**77,866.0 (93.2%)**
State Total	**1,536,340.0**	**21,008.8 (1.4%)**	**19,730.0 (1.3%)**	**16,556.6 (1.1%)**	**48,544.5 (3.2%)**	**1,430,500.1 (93.1%)**
County-level range (low – high)	**71.6 – 50,269.8**	**3.2 – 701.3**	**4.0 - 1,166.9**	**3.4 - 666.8**	**0.0 - 2,134.5**	**37.0 – 49,098.5**

The estimated total length of windbreaks for the six county study area was about 2,345 km while the statewide total was more than 50,000 km (Table 4.8). Complex windbreaks had the longest total length of the four windbreak categories with nearly 19,000 km, which is about 37% of the total length of all

windbreaks in the state. The shortest county-level total length was 12.7 km in Morton County in the southwest corner of the state while Reno County had the longest total windbreak length with 2,172.6 km and was the only county to exceed 2,000 km. The western counties had the shortest total lengths while the longest total lengths were concentrated in the central to south-central portion of the state (Figure 4.8). Overall, there is a general pattern of decreasing county-level windbreak length moving from east to west across the state, with the heaviest concentration of windbreaks in central to south-central part of the state. This high density of windbreaks coincides with the area of Kansas originally targeted for planting windbreaks during the Prairie States Forestry Project (1935-42) mentioned earlier. This area is conducive to growing trees and contains prime farmland but it is also susceptible to severe wind damage (Dahl 1940).

Orientation of single-leg windbreaks also varies by county as evidenced by some counties having longer total lengths of north-south windbreaks versus east-west windbreaks and vice versa. For example, Stafford County has nearly 327 km of east-west oriented windbreaks (53% of total length) compared to only about 63 km (10% of total length) of windbreaks with a north-south orientation (Table 4.8 and refer to Figure 4.7). There are seven counties in which the total length of east-west windbreaks exceeds the total length of north-south windbreaks by more than 80 km and they are concentrated in the central to south-central portion of the state (indicated by the dashed outline in Figure 4.8).

Table 4.8. Length (km) estimates for four windbreak categories derived from a windbreak classification map for a six county study area and the state of Kansas, USA. Percentage of the total length shown in parentheses.

County	Total length of windbreak image objects (km)	Length of image objects classified as north-south windbreak (km)	Length of image objects classified as east-west windbreak (km)	Length of image objects classified as L-shape windbreak (km)	Length of image objects classified as complex windbreak (km)
Chase	343.7	73.0 (21.2%)	61.4 (17.9%)	69.1 (20.1%)	139.9 (40.7%)
Gove	105.7	18.1 (17.1%)	37.6 (35.6%)	34.9 (33.0%)	15.0 (14.2%)
Jackson	576.1	174.9 (30.4%)	72.9 (12.7%)	119.0 (20.7%)	208.9 (36.3%)
Seward	25.2	6.2 (24.6%)	7.1 (28.2%)	7.7 (30.6%)	4.2 (16.7%)
Stafford	613.4	62.7 (10.2%)	326.8 (53.3%)	90.5 (14.8%)	132.9 (21.7%)
Wilson	682.8	182.1 (26.7%)	101.0 (14.8%)	140.9 (20.6%)	258.4 (37.8%)
Study Area Total	**2,345.3**	**517.0 (22.0%)**	**606.8 (25.9%)**	**462.1 (19.7%)**	**759.3 (32.4%)**
State Total	**50,449.3**	**11,198.6 (22.2%)**	**9,418.6 (18.7%)**	**11,105.8 (22.0%)**	**18,726.4 (37.1%)**
County-level range (low – high)	**12.7 – 2,172.6**	**2.4 – 390.1**	**3.1 – 519.2**	**3.2 – 454.6**	**0.0 – 895.4**

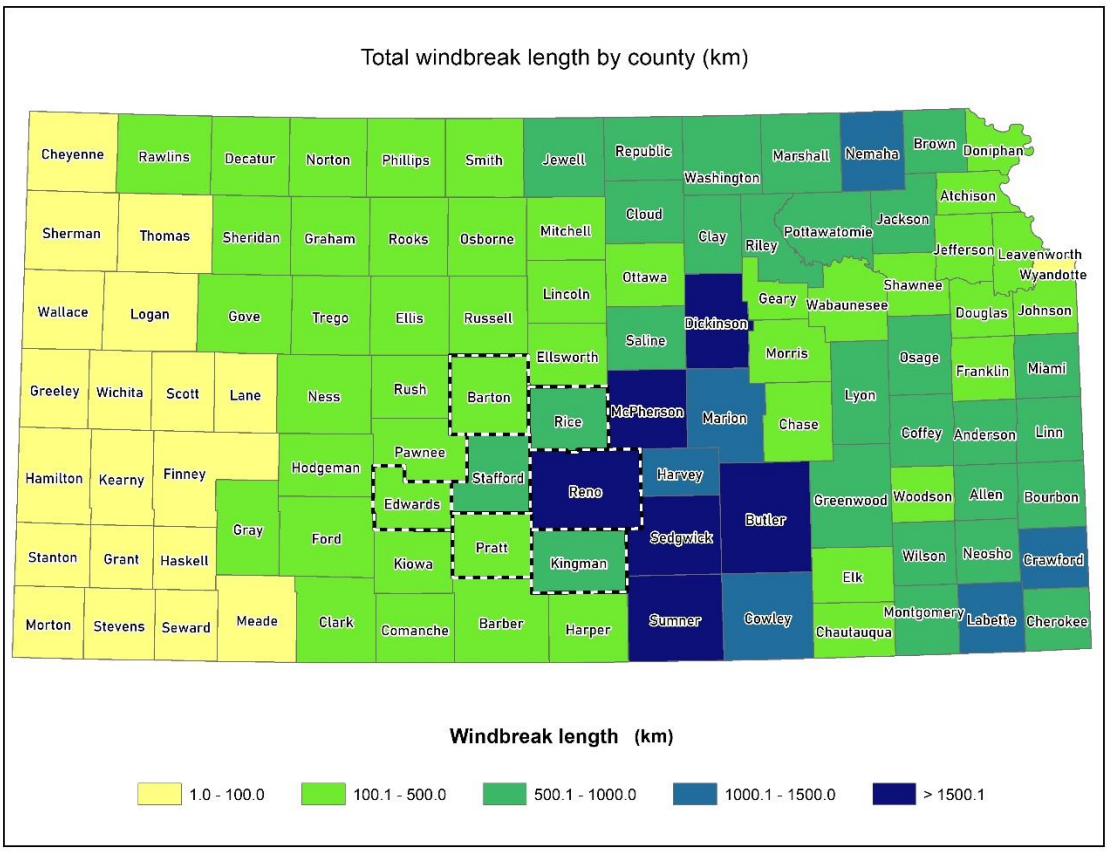


Figure 4.8. Total length of classified windbreaks (km) by county for the state of Kansas, USA. Counties in which the total length of east-west oriented windbreaks exceeds that of north-south windbreaks by more than 80 km indicated by dashed outline.

4.5 Discussion

Our results indicate that a shape-based modelling approach can be used to successfully detect single and multiple-leg windbreaks over large geographic areas with varied densities and patterns of tree cover. Shape indexes by Liknes et al. (2017) proved to be important to the model rather than relying only on the standard shape attributes provided by the eCognition software. The model works especially well for identifying single-leg windbreaks with north-south or east-west orientations as indicated by the high accuracies in both accuracy assessments. However, we observed that there can be confusion between these categories

and the non-windbreak category when these windbreak shapes deviate from linear or are short in length, e.g., less than 100 m in length. The range of configurations for the L-shape and complex windbreak categories also leads to confusion between these classes and also with the non-windbreak category. Generally, much of the misclassification stems from riparian buffers being classified as windbreaks. Incorporating a hydrology layer could alleviate much of this confusion and increase accuracy by distinguishing riparian forest buffers from windbreaks.

Overall accuracy for the five windbreak categories in the six county study area is good with a minimum of 84.1%. The model has some difficulty with consistently separating L-shaped windbreaks from complex windbreaks so combining these classes into one multiple-leg category improves overall accuracy to 87.5% and the user's accuracy for the newly combined class is 80.4%. Combining the first four categories into one windbreak class results in an overall accuracy of 88.8%. Depending on user needs, it may not be necessary to distinguish the windbreaks into separate categories, thus providing a more accurate output product. However, maintaining the orientation information for the single-leg windbreaks may be important. For example, east-west windbreaks are often planted at the south end of fields to protect crops from the hot, desiccating south winds during the summer growing season. Knowing where east-west windbreaks are located, or not located, could be useful for identifying areas that would benefit from planting a windbreak.

The spatial data about windbreaks obtained from this study are important for describing and quantifying the windbreak resource in Kansas. The results tell us how many windbreaks there are as well as their locations, area, length, orientation on the landscape in the case of north-south and east-west windbreaks, and if they are single or have multiple legs. There is little previous statewide information available for us to compare our results. A line intersect sampling study by Hansen (1985) estimated that wooded strips occupied a total area of 136,100 hectares (standard error = 8,700 ha) and totaling 87,000 km

(standard error = 8,400 km) in length. Our results showed smaller area (105,840 ha) and shorter total length (50,450 km) but the aerial photo input data are about three decades apart; thus differences may reflect loss of windbreaks between these two periods. Although limited, there is some research to support the idea that there has been a loss of windbreaks. For example, Baltensperger (1987) found that in Labette County, Kansas, there were 1,087 km of single-leg treed hedgerows based on recent (1966-1980, depending upon availability) aerial photography while our estimate of windbreak length for the county was 425 km based on mapped tree cover from aerial photography acquired in 2015. A similar declining trend was found for two other Kansas counties in his study: Cloud and Coffey. It was estimated that the average hedgerow removal rate was 1.3% for the study area that included areas in Nebraska, Kansas, Iowa, Missouri, and Illinois. This rate of removal would equate to a loss exceeding 20% of the hedgerow resource by the year 2000 (Baltensperger 1987). Bratton et al. (1995) list the number (77,943), hectares (46,113), and kilometers (32,575) of windbreaks in Kansas based on an unpublished report by the Iowa State University Statistical Laboratory. By way of comparison, our results indicate a higher number of windbreaks (118,015), more than twice the area (105,840 ha), and longer total length (50,449 km). This illustrates the need for a consistent method that can be used to establish a baseline inventory of windbreaks from which we can assess changes in the resource over time.

There are some important caveats to consider when using this method, the first being the unknown uncertainty associated with the estimates obtained from the windbreak map. Adequate collection of high-quality reference data would better support the estimates obtained from the output maps, per the recommended good practices in Olofsson et al. (2014). Future work should explore map-based approaches, such as model-based or model-assisted, for estimating uncertainty.

The second caveat to consider when using this method is that the quality of the windbreak output map is related to the quality and accuracy of the input

land cover map. This is because the tree cover image objects are delineated exactly from the tree cover class of the input land cover map; therefore, the quality of the input segments representing groups of trees and windbreaks will impact the classification results. In addition, if windbreaks aren't captured in the original land cover map, then they will not appear in the windbreak map. Data were constrained to the delineated tree cover polygons in the map, meaning interpreters did not manually delineate any windbreaks, so there can be both commission and omission errors that affect the results. The shape of the segments representing windbreaks can be altered by many things. For example, windbreaks with advanced age and/or declining health may have thinned out areas or gaps due to dead or missing trees (e.g., see also Ghimire et al. 2014). In other cases, such as with homestead windbreaks, individual trees in the yard or structures may get lumped in with the windbreak segment, thus altering the overall shape to be other than that of a distinctive windbreak. There are variations of this that can occur with the other types of windbreaks found throughout the various landscapes across the state.

A similar issue, particularly in the eastern portion of the state, is windbreaks connected to larger tracts of forest, mainly riparian buffer systems. The model can't distinguish the windbreak portion from the large segment so the entire segment is labelled as the non-windbreak category, which can lead to an underestimation of total windbreak length. In these aforementioned situations, the result is those segments are labelled as non-windbreaks by the model because the model evaluates the shape of the entire segment; perhaps a next-generation model could separate the windbreak portion from the larger tract of forest. However, these issues are an artifact of the land cover map creation process and are beyond our control, with the possible exception of how windbreaks attached to forest patches are classified. As previously mentioned, improved analysis techniques may be able to identify these attached windbreaks in future assessments. These and other misclassification issues can be corrected with manual editing. Tracking the corrections made could quantify the impact of

making these changes, i.e., how much additional area and length of windbreaks is gained by including these previously excluded windbreaks. This is important because of the specific interest in windbreaks and quantifying the ecosystem services they provide.

Another challenge we encountered is that windbreaks of all categories come in a variety of shapes and configurations, some of which deviate from the regular, straight-line shape of the windbreak leg(s). This occurs most frequently in the complex category. These complex systems with multiple legs can also be labelled as non-windbreaks by the model. We attempted to include such variations in all categories in our training data but found that including these types of anomalies yielded poorer classification results. Selecting high-quality, representative training data samples with clear, distinctive shapes from each windbreak category is key to obtaining good classification results. This is likely one reason why the OOB accuracy (refer to Table 4.4) was higher than the photo-interpreted accuracy (refer to Table 4.5). The OOB sample is made up of the training data, which are carefully selected to obtain high-quality samples whereas the samples in the independent accuracy assessment are selected at random so there is higher variability in sample shape and quality.

Even with good initial results, we still recommend making manual corrections. This takes time but is much faster than potentially more than a year's worth of time required to manually delineate each individual windbreak (e.g., Piwowar et al. 2017). Making corrections is generally considered part of the workflow when working with high resolution data (e.g., O'Neil-Dunne et al. 2014; Ahles et al. 2016) and we concur with recommendations to include this step. Post-modeling manual editing serves as a data quality assurance procedure and ultimately results in an improved output product that most likely has higher accuracy. Furthermore, the incorporation of aerial photography provides context information that is not available when examining shape alone. This allows the interpreter to deduce that a group of trees is functioning as a windbreak even if the shape is unlike a traditional windbreak, thus enhancing the inventory. This

may partially explain why the human interpreted accuracy (refer to Table 4.5) is lower than the OOB accuracy (refer to Table 4.4). The shape of a group of trees may not have the distinct shape of one of the four windbreak categories due to factors mention above, but human cognition can determine that the group of trees is serving a windbreak function based on what is seen in the imagery. When manual editing is completed, the statewide dataset of Kansas windbreaks will be published to the USDA Forest Service Research Data Archive (https://www.fs.usda.gov/rds/archive/) as a collection of county-level windbreak datasets available for download at no cost to the user.

4.6 Conclusions

The objective of this study was to develop a practical, operational method for identifying both single and multiple-leg treed windbreaks in high-resolution (1m) land cover maps for large geographic regions. To that end, we developed a GEOBIA modeling approach using standard segmentation procedures coupled with an improved supervised classification approach to identify four windbreak categories as well as a non-windbreak category across a range of extents and patterns of tree cover in Kansas, USA, using a 1 m publicly available land cover dataset. This approach overcomes two shortcomings noted in the literature pertaining to windbreak mapping studies. First, it addresses the lack of semi-automated windbreak detection methods implemented for large geographic areas as noted by Pasher et al. (2016) and Vannier and Hubert-Moy (2014), and second, it represents an incremental improvement to the identification of more complex multiple-leg windbreaks, particularly L-shaped windbreaks, a difficulty noted by Liknes et al. (2017) and Ha et al. (2019).

This approach stands in contrast to methods that aim to detect windbreak objects directly in imagery. We have chosen to focus on extracting these landscape features from existing high spatial resolution land cover datasets because those data are becoming more available in many parts of the world. This is an important step in large-scale operational mapping of windbreaks and is

the first time that a detailed and spatially explicit windbreak dataset has been produced for the entire state of Kansas, USA. Such data are critical for value-added analyses that will quantify the important ecological and economic services provided by windbreaks.

4.7 References

Ahles, N., MacFaden, S., O'Neil-Dunne, J., Royar, A., Engel, T. 2016. GEOBIA systems for massive data processing. Available online: https://doi.org/10.3990/2.457 (accessed on 17 November 2019)

Aksoy, S., Akçay, H.G., Wassenaar, T., 2010. Automatic mapping of linear woody vegetation features in agricultural landscapes using very high resolution imagery. IEEE Trans. Geosci. Remote Sens. 48 (1), 511-522.

Amichev, B.Y., Bentham, M.J., Cerkowniak, D., Kort, J., Kulshreshtha, S., Laroque, C.P., Piwowar, J.M., Van Rees, K.C.J., 2015. Mapping and quantification of planted tree and shrub shelterbelts in Saskatchewan, Canada. Agroforest. Syst. 89, 49-65.

Ballesteros-Possu, W., Brandle, J., Schoeneberger, M., 2017. Potential of windbreak trees to reduce carbon emissions by agricultural operations in the US. Forests 8(5), 138; https://doi.org/10.3390/f8050138

Baltensperger, B.H., 1987. Hedgerow distribution and removal in nonforested regions of the Midwest. J. Soil Water Cons. 42(1), 60-64.

Basu, S., Ganguly, S., Nemani, R.R., Mukhopadhyay, S., Zhang, G., Milesi, C., Michaelis, A., Votava, P., Dubayah, R., Duncanson, L., et al., 2015. A semiautomated probabilistic framework for tree-cover delineation from 1-m NAIP imagery using a high-performance computing architecture. IEEE Trans. Geosci. Remote Sens. 53, 5690–5708.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens. 114, 24-31.

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M., 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS J. Photogramm. Remote Sens. 58, 239-258.

Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS J. Photogramm. 65, 2-16.

Brandle, J., Finch, S., 1991. EC91-1763-B How windbreaks work. Historical Materials from University of Nebraska-Lincoln Extension. 4709.

Http://digitalcommons.unl.edu/extenstionhist/4709 last accessed 2/23/2019.

Brandle, J.R., Hodges, L., Zhou, X.H., 2004. Windbreaks in North American Agricultural Systems. Agroforest. Syst. 61, 65-78.

Bratton, G.F., Schaefer, P.R., Brandle, J.R., Conservation Forestry for Sustainable Great Plains Ecosystems in book: Conservation of Great Plains Ecosystems, Kluwer Academic Publishers, 211-227.

Breiman, L., 2001a. Random forests. Mach. Learn. 45, 5-32.

Breiman, L., Cutler, A., (n.d.). Random Forests. Available online: https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm#papers (accessed on 10 March 2020).

Chesapeake Conservancy, 2016. Land Cover Data Project 2013/2014. Available online: https://chesapeakeconservancy.org/conservation-innovation-center/land-cover-data-project/ (accessed on July 8, 2020).

Congalton, R.G., Green, K., 2009. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 2nd ed.; Taylor & Francis Group, LLC: Boca Raton, FL, USA, pp. 63-83.

Cutler, D., Edwards, Jr., T., Beard, K., Cutler, A., Hess, K., Gibson, J., Lawler, J., 2007. Random forests for classification in ecology. Ecology 88, 2783-2792.

Czerepowicz, L., Case, B.S., Doscher, C. Using satellite image data to estimate aboveground shelterbelt carbon stocks across an agricultural landscape. Agric., Ecosyst. Environ. 2012, 156, 142-150.

Dahl, J., 1940. Progress and development of the Prairie States Forestry Project. J. For. 38, 301-306.

Dronen, S.I., 1984. Windbreaks in the Great Plains. North. J. Appl. For. 3, 55-59.

Ghimire, K., Dulin, M.W., Atchison, R.L., Goodin, D.G., Hutchinson, J.M.S., 2014. Identification of windbreaks in Kansas using object-based image analysis, GIS techniques and field survey. Agroforest. Syst. 88, 865-875.

Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random Forests for landcover classification. Pattern Recognit. Remote Sens. 27, 294-300.

Ha, T.V., Amichev, B.Y., Belcher, K.W., Bentham, M.J., Kulshreshtha, S.N., Laroque, C.P., Van Rees, K.C.J., 2019. Shelterbelt Agroforestry Systems Inventory and Removal Analyzed by Object-based Classification of Satellite Data in Saskatchewan, Canada. Can. J. Remote Sens. 45, 246-263.

Hansen, M.H., 1985. Line intersect sampling of wooded strips. Forest Sci. 31 (2), 282-288.

Hart, R.H., Hart, J.A., 1997. Rangelands of the Great Plains before European Settlement. Rangelands, 19 (1), 4-11.

Hartong, A.L., Moessner, K.E., 1956. Wooded strips in Iowa. Forest Survey Release 21. USDA Central States Forest Experiment Station, Columbus, OH.

Kellerman, T.A., Meneguzzo, D.M., Vaitkus, M., White, M., Ossell, R., Sorsen, N., Stannard, J., Gift, T., Cox, J., Liknes, G.C., 2019. High-resolution land cover of Nebraska (2014). Fort Collins, CO: Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2019-0038.

Land Cover Data Project. Available online: https://chesapeakeconservancy.org/conservation-innovation-center-2/high-resolution-data/land-cover-data-project/ (accessed on 9 March 2020).

Liknes, G.C., Perry, C.H., Meneguzzo, D.M., 2010. Assessing tree cover in agricultural landscapes using high-resolution aerial imagery. J. Terr. Obs. 2(1): 38-55.

Liknes, G.C., Meneguzzo, D.M., Kellerman, T.A., 2017. Shape indexes for semi-automated detection of windbreaks in thematic tree cover maps from the central United States. Int. J. Appl. Earth Obs. Geo-Inf. 59, 167-174.

Lunetta, R. S., 2004. Remote Sensing and GIS Accuracy Assessment. United Kingdom: CRC Press.

Martínez, A.M., Kak, A.C., 2001. PCA versus LDA. IEEE T. Pattern Anal. 23 (2), 228-233.

Maxwell, A.E., Strager, M.P., Warner, T.A., Ramezan, C.A., Morgan, A.N., Pauley, C.E., 2019. Large-Area, High Spatial Resolution Land Cover Mapping Using Random Forests, GEOBIA, and NAIP Orthophotography: Findings and Recommendations. Remote Sens. 11, 1409.

Meneguzzo, Dacia M.; Liknes, Greg C.; Nelson, Mark D. 2013. Mapping trees outside forests using high-resolution aerial imagery: a comparison of pixel- and object based classification approaches. Environ. Monit. Assess. 185: 6261-6275.

Meneguzzo, D.M., Lister, A.J., Sullivan, C., 2018. Summary of findings from the Great Plains Tree and Forest Invasives Initiative. Gen. Tech. Rep. NRS-GTR-177. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. 24 p. https://doi.org/10.2737/NRS-GTR-177.

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42-57.

Olson, C.E., 1960. Elements of photographic interpretation common to several sensors. Photogramm. Eng. 26b (4), 651–656.

O'Neil-Dunne, J., Pelletier, K., MacFaden, S., Troy, A., Grove, J.M., 2009. Object-based high-resolution land-cover mapping: operational considerations. In 2009 17th International Conference on Geoinformatics (pp. 1-6). IEEE.

O'Neil-Dunne, J.O., MacFaden, S., Royar, A., 2014. A Versatile, Production-Oriented Approach to High-Resolution Tree-Canopy Mapping in Urban and Suburban Landscapes Using GEOBIA and Data Fusion. Remote Sens. 6, 12837-12865.

Pankiw, J., 2013. The use of object-based classification of high resolution panchromatic satellite imagery for the inventory of shelterbelts in the province of Saskatchewan, Faculty of Graduate Studies and Research. SK, Canada: University of Regina.

Pankiw, J., Piwowar, J.M., 2010. Seasonality of imagery: the impact on object-based classification accuracy of shelterbelts. Prairie Perspectives: Geographical Essays, 13, 39-48.

Pariona, A., 2017. Least Densely Populated U.S. States. Available online: worldatlas.com/articles/least-densely-populated-u-s-states.html (accessed on 7 February 2020).

Pasher, J., McGovern, M., Putinski, V., 2016. Measuring and monitoring linear woody features in agricultural landscapes through earth observation data as an indicator of habitat availability. Int. J. Appl. Earth Obs. Geo-Inf. 44, 113-123.

Paull, D.A., Whitson, J.W., Marcotte, A.L., Liknes, G.C., Meneguzzo, D.M., Kellerman, T.A., 2017. High-resolution land cover of Kansas (2015). Available online: https://doi.org/10.2737/RDS-2017-0025

PBS. "Surviving the Dust Bowl: The Drought." http://www.pbs.org. Available online: https://www.pbs.org/wgbh/americanexperience/features/dustbowl-drought/ (Accessed July 5, 2020).

Perry, C.H., Woodall, C.W., Liknes, G.C., Schoeneberger, M.M., 2009. Filling the gap: improving estimates of working tree resources in agricultural landscapes. Agroforest. Syst. 75, 91-101.

Peterson, D.L., Egbert, S.L., Price, K.P., Martinko, E.A., 2004. Identifying historical and recent land-cover changes in Kansas using post-

classification change detection techniques. T. Kans. Acad. Sci. 107, 105-118.

Piwowar, J.M., Amichev, B.Y., Van Rees, K.C.J., 2017. The Saskatchewan shelterbelt inventory. Can. J. Soil Sci. 97, 433-438.

Ramachandran Nair, P.K., Mohan Kumar, B., Nair, V.D., 2009. Agroforestry as a strategy for carbon sequestration. J. Plant Nutr. Soil Sci. 172 (1), 10-23.

Read, R., 1958. The Great Plains shelterbelt in 1954: a re-evaluation of field windbreaks planted between 1935 and 1942 and a suggested research program. Great Plains Agric. Counc. No. 16.

Robinson, C., Hou, L., Malkin, K., Soobitsky, R., Czawlytko, J., Dilkina, B., Jojic, N. Large Scale High-Resolution Land Cover Mapping with Multi-Resolution Data. Proceedings of the 2019 Conference on Computer Vision and Pattern Recognition (CVPR 2019).

Sauer, T.J., 2009. The Prairie States Forestry Project as a Model for an Effective Global Climate Change Mitigation Project. In: Handbook of Agroforestry Management Practices, L.R. Kellimore (ed.), Nova Science Publishers, Inc.

Schmidt, T.L., Wardle, T.D., 1998. The forest resources of Nebraska. Res. Pap. NC-332. St. Paul, MN: U.S. Dept. of Agriculture, Forest Service, North Central Research Station (https://doi.org/10.2737/NC-RP-332)

Schoeneberger, M.M., Bentrup, G., de Gooijer, H., Soolanayakanahally, R., Sauer, T., Brandle, J., Zhou, X., Current, D., 2012. Branching out: agroforestry as a climate change mitigation and adaptation tool for agriculture. J. Soil Water Conserv. 67 (5), 128A-136A.

St. Peter, J., Hogland, J., Anderson, N., Drake, J., Medley, P., 2018. Fine Resolution Probabilistic Land Cover Classification of Landscapes in the Southeastern United States. ISPRS Int. J. Geo-Inf. 7, 107.

“The Establishment of a Forest Shelter Belt.” 1934. Science, 80 (2065), p. 91. DOI: 10.1126/science.80.2065.91

van der Werff, H.M.A., van der Meer, F.D., 2008. Shape-based classification of spectrally identical objects. ISPRS J. Photogramm. Remote Sens. 63 (2), 251–258.

Vannier, C. and L. Hubert-Moy. 2014. Multiscale comparison of remote sensing data for linear woody feature vegetation mapping. Int. J. Remote Sens. 35(21), 7376-7399.

Wiseman, G., Kort, J., Walker, D., 2009. Quantification of shelterbelt characteristics using high-resolution imagery. Agr. Ecosyst. Environ. 131, 111-117.

USDA Forest Service, 1937. Forestry for the Great Plains. Washington, D.C.: Forest Service, U.S. Dept. of Agriculture.

USDA Forest Service, Forest Inventory and Analysis Program, Thu Jul 09 15:25:08 GMT 2020. Forest Inventory EVALIDator web-application Version 1.8.0.01. St. Paul, MN: U.S. Department of Agriculture, Forest Service, Northern Research Station. [Available only on internet: http://apps.fs.usda.gov/Evalidator/evalidator.jsp]

USDA National Agricultural Statistics Service CropScape - Cropland Data Layer. Available online: https://nassgeodata.gmu.edu/CropScape/ (accessed on 6 February 2020).

USDA National Agroforestry Center, 2012. What is a windbreak? Available online: https://www.fs.usda.gov/nac/assets/documents/workingtrees/infosheets/wb_info_050712v8.pdf (accessed on July 8, 2020).

Yan, L., Roy, D.P., 2015. Improved time series land cover classification by missing-observation-adaptive nonlinear dimensionality reduction. Remote Sens. Environ. 158, 478-491.

Chapter 5. Conclusions

The overall goal of this research was to develop remote sensing-based approaches to answer three questions regarding tree resources in the Great Plains region of the United States: how much tree cover is out there? Where is it? What is it doing? Three separate studies were conducted as a means to provide answers to these questions and are presented in Chapters 2-4 of this dissertation. They describe the development of robust and operational methods for mapping tree cover from high-resolution aerial photography from the National Agriculture Imagery Program (NAIP) and then subsequently classifying the tree cover class into one of four windbreak categories or a non-windbreak category. These semi-automated approaches can be used to conduct broad-scale comprehensive assessments of tree cover and one very important tree functional category: windbreaks. They produce detailed spatial information about each group of trees that can be summarized to describe the tree cover resource at a range of scales: from individual landholdings to county, state, or regional levels.

In the first study (Chapter 2), two remote sensing-based approaches were identified that use publicly available high-resolution (1 m) NAIP imagery to map all tree cover in an agriculturally-dominant landscape. The results obtained using an unsupervised per-pixel classifier (independent component analysis – [ICA]) and an object-based image analysis (OBIA) procedure in Steele County, Minnesota, USA, were compared using three types of accuracy assessments. These assessments evaluated how each method performed in terms of: 1) producing a county-level estimate of total tree-covered area, 2) correctly locating tree cover on the ground, and 3) how tree cover patch metrics computed from the output compared to those delineated by a human photo interpreter. Both methods produced reliable maps of tree cover versus no-tree cover over a broad spatial extent and could serve to supplement ground-based inventory data. The OBIA approach had higher overall accuracy (95%) compared to the ICA approach (88%). The ICA approach produced an estimate of total tree cover

more similar to the photo-interpreted result, but the output from the OBIA method was more accurate in terms of describing the actual observed spatial pattern of tree cover and produced a more realistic output product. Because the OBIA approach produced classification results that were more accurate in terms of spatial location and also provides more reasonable information about the spatial pattern of tree cover, it is the better choice for mapping tree cover. This study also concluded that urban landscapes are extremely complex and it is difficult to accurately delineate tree cover using either of the remote sensing-based approaches so a different classification model needs to be developed for urban landscapes. As a result, future work will focus on rural landscapes only.

The goal of Chapter 3 was to develop a robust and transferable method for operational mapping of general categories of land cover in rural landscapes using 1 m NAIP imagery as the sole data source for a four-state study area (Kansas, Nebraska, North Dakota, and South Dakota) in the Great Plains region of the central United States. Building on research from Chapter 2, improvements were made to the previous OBIA method, including: 1) more semi- and automated processing to make the approach more efficient and operational, 2) additional land cover classes, 3) building a segmentation routine that works generally well across a variety of landscapes so settings do not have to be adjusted within the four-state study area, and 4) the development of a more robust classification approach to accommodate the spectral and illumination issues associated with NAIP imagery. The updated method was successfully applied to all four states and land cover maps for Kansas and Nebraska, USA, have been published and are available for download on the United States Department of Agriculture Research Data Archive website (https://www.fs.usda.gov/rds/archive/). Maps for North Dakota and South Dakota are forthcoming. The maps were found to accurately represent tree canopies, with accuracies of 98.6 and 98.5% at the state level for Kansas and Nebraska, respectively, thus providing data at an appropriate scale for land managers in this region.

There are several aspects of this study that distinguish it from other NAIP-based classification studies, especially those conducted for large geographic areas. The first is that no ancillary data were utilized in the segmentation or classification processes. The second is that a single segmentation routine was developed and applied within the four-state study area; no adjustments were made when moving from state to state. Thirdly, the developed segmentation algorithm utilized information in the green spectral band and one of its derivatives: an edge detection layer created by applying a 3x3 sobel operator filter to form the image objects. This is different from other statewide land cover mapping studies that used all four spectral bands of NAIP imagery weighted equally for segmentation.

In the last study (Chapter 4), the output dataset produced for Kansas via the method developed in Chapter 3 serves as the input data source when answering the final question behind this research: what is the tree cover doing? This question can be partially answered by identifying a prominent tree functional category: windbreaks. Windbreaks are recognizable because of their distinctive linear shapes. The objective of Chapter 4 was to develop a practical, operational method for identifying both single and multiple-leg treed windbreaks from high-resolution (1 m) land cover maps that would be scalable for large regions with varied densities and patterns of tree cover. It was hypothesized that single and multiple-leg windbreaks could be distinguished from other groups of tree cover based on shape-related information alone. A shape-based classification model was developed and applied successfully to the Kansas land cover dataset, and more than 50,000 km of windbreaks occupying 105,840 hectares were identified. An overall accuracy of 84.1% was achieved for four different windbreak classes and a non-windbreak category based on an accuracy assessment by an independent interpreter; overall accuracy increases to 88.8% after aggregating the windbreak classes into a single windbreak category versus a non-windbreak category. This research overcomes two deficiencies noted in the literature pertaining to windbreak mapping studies: 1) it addresses the lack of semi-

automated windbreak detection methods implemented for large geographic areas, and 2) it represents an incremental improvement to the identification of more complex multiple-leg windbreaks, particularly L-shaped windbreaks. This approach stands in contrast to the many studies that aim to detect windbreak objects directly in imagery. Rather, it focuses on extracting these landscape features from existing high spatial resolution land cover datasets because those data are becoming more readily available in many parts of the world. This is an important step in large-scale operational mapping of windbreaks and is the first time that a detailed and spatially explicit windbreak dataset has been produced for the entire state of Kansas, USA. Such data are critical for value-added analyses that will quantify the important ecological and economic benefits provided by windbreaks.

The implications of this research are that comprehensive assessments of tree cover and windbreaks in the Great Plains region can be conducted via the methods presented in this dissertation. The developed methods are applicable over large geographic areas and the output can be scaled up from individual landholdings to the state or even regional levels, which is important for consistent broad-scale inventory and monitoring of tree resources over time. In addition, having readily available fine-scale geospatial data products at these levels is important for natural resource professionals and decision makers. This endeavor is the first of its kind and provides information at a new scale that is appropriate for inventory, monitoring, and decision making regarding tree resources in the Great Plains.

Complete Bibliography

Ahles, N., MacFaden, S., O'Neil-Dunne, J., Royar, A., & Engel, T. (2016) GEOBIA systems for massive data processing. Retrieved from https://doi.org/10.3990/2.457

Aksoy, S., Akcay, H. G., & Wassenaar, T. (2010). Automatic mapping of linear woody vegetation features in agricultural landscapes using very high resolution imagery. *IEEE Transactions on Geoscience and Remote Sensing, 48*(1), 511-522.

Amichev, B.Y., Bentham, M.J., Cerkowniak, D., Kort, J., Kulshreshtha, S., Laroque, C.P., Piwowar, J.M., & Van Rees, K.C.J. (2015). Mapping and quantification of planted tree and shrub shelterbelts in Saskatchewan, Canada. *Agroforestry Systems 89*, 49-65.

Amorós-López, J., Izquierdo-Verdiguier, E., Gomez-Chova, L., Munoz-Mari, J., Rodriguez-Barreiro, J.Z., Camps-Valls, G., & Calpe-Maravilla, J. (2011). Land cover classification of VHR airborne images for citrus grove identification. *ISPRS Journal of Photogrammetry and Remote Sensing, 66,* 115-123.

Anderson, J.R., Hardy, E.E., Roach J.T., & Witmer, R.E. (1976). A land use and land cover classification system for use with remote sensor data. U.S. Geological Survey Professional Paper 964. Washington D.C.

Andrews, H.J., & Bromley, W.S. (1941). *Trends in land use in northern Michigan: A study of Alpena, Antrim, Ogemaw and Roscommon Counties*. The Charles Lathrop Park Forestry Foundation, Washington D.C.

Baatz, M., & Schape, M. (2000). Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation. In J. Strobl, T. Blaschke, & G. Griesebner (Eds.), *Angewandte Geographische Informations-Verarbeitung XII* (pp. 12-23). Wichmann Verlag, Karlsruhe.

Ballesteros-Possu, W., Brandle, J., & Schoeneberger, M. (2017). Potential of windbreak trees to reduce carbon emissions by agricultural operations in the US. *Forests 8*(5), 138. doi:10.3390/f8050138

Baltensperger, B.H. (1987). Hedgerow distribution and removal in nonforested regions of the Midwest. *Journal of Soil and Water Conservation 42*(1), 60-64.

Basu et al. (2015). A semiautomated probabilistic framework for tree-cover delineation from 1-m NAIP imagery using a high-performance computing architecture. *IEEE Transactions on Geoscience and Remote Sensing, 53*(10), 5690-5708. doi: 10.1109/TGRS.2015.2428197

Bauer, M.E., Burk, T.E., Ek, A.R., Coppin, P.R., Lime, S.D., Walsh, T.A., Walters, D.K., Befort, W., & Heinzen, D.F. (1994). Satellite inventory of Minnesota Forest Resources. *Photogrammetric Engineering & Remote Sensing, 60*(3), 287-298.

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing, 114*, 24-31.

Bellefontaine, R., Petit, S., Pain-Orcet, M., Deleporte, P., & Bertault, J.G. (2002). Trees outside forests. Towards better awareness. Retrieved from http://www.fao.org/3/y2328e/y2328e00.htm

Bentrup, G., Hopwood, J., Adamson, N., & Vaughan, M. (2019). Temperate Agroforestry Systems and Insect Pollinators: A Review. *Forests, 10*, 981. doi: 10.3390/f10110981

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing, 58*(3-4), 239-258.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing 65*, 2-16.

Blaschke, T., & Strobl, J. (2001). What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. *GIS – Zeitschrift fur Geoinformationssysteme, 14*(6), 12-17.

Blaschke, T., & Hay, G.J. (2001). Object-oriented image analysis and scale-space theory: Theory and methods for modeling and evaluating multi-scale landscape structure. *International Archives of Photogrammetry and Remote Sensing, 354*, 22-29.

Blaschke, T., & Lang, S. (2006). Object based image analysis for automated information extraction – A synthesis. In Measuring the Earth II APSRS Fall Conference, San Antonio, TX, USA, 6-10 November 2006 (pp. 6-10).

Blaschke, T., Lang, S., Lorup, E., Strobl, J., & Zeil, P. (2000). Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. In A. Cremers, & K. Greve (Eds.), *Environmental Information for Planning, Politics and the Public*, vol. 2 (pp. 555-570). Marburg: Metropolis Verlag.

Blaschke, T., Burnett, C., & Pekkarinen, A. (2004). New contextual approaches using image segmentation for object-based classification. In F. De Meer, &S. de Jong (Eds.), *Remote Sensing Image Analysis: Including the spatial domain* (pp. 211-236). Dordrecht: Kluver Academic Publishers.

Blaschke, T., Lang, S., & Hay, G.J. (Eds.). (2008). *Object Based Image Analysis*. New York: Springer.

Brandle, J., & Finch, S. (1991). How windbreaks work. Retrieved from http://digitalcommons.unl.edu/extenstionhist/4709

Brandle, J.R., Hodges, L., & Zhou, X.H. (2004). Windbreaks in North American Agricultural Systems. *Agroforestry Systems, 61*, 65-78.

Brandle, J.R., Wardle, T.D., & Bratton G.F. (1992). Opportunities to increase tree planting in shelterbelts and the potential impacts on carbon storage and conservation. In N.R. Sampson, & F. Hair F (Eds.), *Forests and Global Change, Vol. 1. Opportunities for Increasing Forest Cover* (pp. 157-176). Washington D.C.: American Forests.

Bratton, G.F., Schaefer, P.R., & Brandle, J.R. (1997). Conservation Forestry for Sustainable Great Plains Ecosystems. In S.R. Johnson, & A. Bouzaher (Eds), *Conservation of Great Plains Ecosystems* (pp. 211-227). Kluwer Academic Publishers.

Breiman, L. (1994). Bagging predictors. Retrieved from https://statistics.berkeley.edu/sites/default/files/tech-reports/421.pdf

Breiman, L. (2001a). Random forests. *Machine Learning, 45*, 5-32.

Breiman, L. (2001b). Statistical modeling: the two cultures. *Statistical Science, 16*(3), 199-231.

Breiman, L., & Cutler, A. (n.d.) Random Forests. Retrieved from https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm#papers

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and Regression Trees*. Belmont, CA: Wadsworth International Group.

Briem, G.J., Benediktsson, J.A., & Sveinsson, J.R. (2002). Multiple classifiers applied to multisource remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing, 40*, 2291–2299.

Bryant, J. (1979). On the clustering of multidimensional pictorial data. *Pattern Recognition, 11*, 115-125.

Burnett, C., & Blaschke, T. (2003). A multi-scale segmentation/object relationship modeling methodology for landscape analysis. *Ecological Modeling, 168*(3), 233-249.

Campbell, J. (1996). *Introduction to Remote Sensing* (2nd ed.) New York, NY: The Guilford Press.

Castaneda, O., Chhin, S., Chomentowski, W., Jindal, R., Justice, C., Kasten, E., Loboda, T., MacFarlene, D., Samek, J., Skole, D., & Smalligan, D. (2011). Guidelines for Measuring Carbon in Trees Outside Forests. Retrieved from http://www.goes.msu.edu/cbp/Module2.pdf

Castillejo-González, I. L., López-Granados, F., García-Ferrer, A., Peña-Barragán, J. M., Jurado-Expósito, M., de la Orden, M. S, et al. (2009). Object- and pixel-based analysis for mapping crops and their agro-environmental associated measures using QuickBird imagery. *Computers and Electronics in Agriculture, 68*(2), 207–215.

Chan, J., & Paelinckx, D. (2008). Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotype mapping using airborne hyperspectral imagery. *Remote Sensing of Environment, 112*(6), 2999–3011.

Chen, F., Guan, Z., Yang, X., & Cui, W. (2011). A novel remote sensing image fusion method based on independent component analysis. *International Journal of Remote Sensing, 32*(10), 2745-2763.

Chen, G., Hay, G.J., Carvalho, L.M., & Wulder, M.A. (2012). Object-based change detection. *International Journal of Remote Sensing, 33*(14), 4434-4457.

Chubey, M. S., Franklin, S. E., & Wulder, M. A. (2006). Object-based Analysis of Ikonos-2 Imagery for Extraction of Forest Inventory Parameters. *Photogrammetric Engineering & Remote Sensing, 72*(4), 383-394.

Chesapeake Conservancy. (2016). Land Cover Data Project. Retrieved from https://chesapeakeconservancy.org/conservation-innovation-center-2/high-resolution-data/land-cover-data-project/

Cibula, W.G., & Nyquist, M.O. (1987). Use of topographic and climatological models in a geographic data base to improve Landsat MSS classification for Olympic National Park. *Photogrammetric Engineering and Remote Sensing, 53*, 67-75.

Cihlar, J., Ly, H., & Xiao, Q. (1996). Land cover classification with AVHRR multichannel composites in northern environments. *Remote Sensing of Environment, 58*, 36-51.

Cleve, C., Kelly, M., Kearns, F.R., & Moritz, M. (2008). Classification of the wildland-urban interface: A comparison of pixel- and object-based

classifications using high-resolution aerial photography. *Computers, Environment and Urban Systems 32*, 317-326.

Cohen, W.B., Kushla, J.D., Ripple, W.J., & Garman, S.L. (1996). An introduction to digital methods in remote sensing of forested ecosystems: focus on the Pacific Northwest, USA. *Environmental Management, 20*(3), 421-435.

Cohen, W.B., Spies, T.A., & Fiorella, M. (1995). Estimating the age and structure of forests in a multiownership landscape of western Oregon, USA. *International Journal of Remote Sensing, 16*, 721-746.

Colwell, R.N. (1956). Determining the prevalence of certain cereal crop diseases by means of aerial color photography. *Hilgardia, 26*(5), 223-286.

Colwell, R.N. (1965). The extraction of data from aerial photographs by human and mechanical means. *Photogrammetria*, *20*, 211-228.

Colwell, R.N. (1968). The field of photo interpretation. Pages 1-15 in Proc. 2nd Semin. Aerial Photointepret. Dev. Can. Queen's Printer and Controller of Stationery, Ottawa.

Comon, P. (1994). Independent component analysis, a new concept? *Signal Processing 36*(3), 287-314.

Congalton, R.G. (1991). A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote Sensing of Environment, 37*, 35-46.

Congalton, R.G. 2010. Remote sensing: an overview. GIScience & Remote Sensing, 47(4), 443-459.

Congalton, R.G., Birch, K., Jones, R., & Schriever, J. 2002. Evaluating remotely sensed techniques for mapping riparian vegetation. *Computers and Electronics in Agriculture 37*, 113-126.

Congalton, R.G., & Green, K. (2009). Assessing the Accuracy of Remotely Sensed Data: Principles and Practices (2nd ed.)(pp. 63-83). Boca Raton, FL: Taylor & Francis Group, LLC.

Congalton, R.G., Green, K., & Teply, J. (1993). Mapping old growth forests on national forest and park lands in the Pacific Northwest from remotely sensed data. *Photogrammetric Engineering and Remote Sensing, 59,* 529-535.

Coppin, P.R. (1991). The change component in multitemporal Landsat TM Images: its potential for forest inventory and management. PhD thesis, University of Minnesota, St. Paul.

Coulston, J. W., Moisen, G.G., Wilson, B.T., Finco, M.V., Cohen, W.B., & Brewer, C.K. (2012). Modeling percent tree canopy cover: a pilot study. *Photogrammetric Engineering and Remote Sensing, 78,* 715–727.

Curran, P.J. (1985). *Principles of Remote Sensing*. London: Longman Publishing Group.

Culvenor, D. (2003). Extracting individual tree information: a survey of techniques for high spatial resolution imagery. In M. Wulder, & S. Franklin (Eds.), *Remote Sensing of Forest Environments: Concepts and Case Studies* (p. 255-277). Boston/Dordecht/London: Kluwer Academic Publishers.

Cutler, D., Edwards, Jr., T., Beard, K., Cutler, A., Hess, K, Gibson, J. & Lawler, J. (2007). Random forests for classification in ecology. *Ecology, 88*(11), 2783-2792.

Czerepowicz, L., Case, B.S., & Doscher, C. (2012). Using satellite image data to estimate aboveground shelterbelt carbon stocks across an agricultural landscape. *Agriculture, Ecosystems & Environment, 156*, 142-150.

Dahl, J. (1940). Progress and development of the Prairie States Forestry Project. *Journal of Forestry, 38*, 301-306.

Davies, K. W., Petersen, S. L., Johnson, D. D., Davis, D. B., Madsen, M. D., Zvirzdin, D. L., & Bates, J. D. (2010). Estimating juniper cover from National Agriculture Imagery Program (NAIP) imagery and evaluating relationships between potential cover and environmental variables. *Rangeland Ecology & Management, 63*(6), 630-637.

Definiens AG. (2010). eCognition Developer 8.0.1 Reference Book. Retrieved from http://www.definiens.com.

de Foresta, H., Somarriba, E., Temu, A., Boulanger, D., Feuilly, H., & Gauthier, M., 2013. Towards the Assessment of Trees Outside Forests. Retrieved from http://www.fao.org/3/aq071e/aq071e00.htm

DeFries, R., Hansen, M., & Townshend, J. (1995). Global discrimination of land cover types from metrics derived from AVHRR pathfinder data. *Remote Sensing of Environment, 54*, 209-222.

DeFries, R.S., Hansen, M., & Townshend, J.R.G. (1998). Global land cover classifications at 8 km spatial resolution: The use of training data derived from Landsat imagery in decision tree classifiers. *International Journal of Remote Sensing, 19*(16), 3141-3168.

De Jong, S.M., & van der Meer, F.D. (Eds.) (2004). *Remote Sensing Image Analysis: Including the Spatial Domain,* vol. 5. Dordrecht: Kluwer Academic Publishers.

Dingle Robertson, L., & King, D. J. (2011). Comparison of pixel- and object-based classification in land cover change mapping. *International Journal of Remote Sensing, 32*(6), 1505–1529.

Dorren, K.A., Maier, B., & Seijmonsbergen, A.C. (2003). Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification. *Forest Ecology and Management, 183*, 31-46.

Drăgut, L., & T. Blaschke. (2006). Automated classification of landform elements using object-based image analysis. *Geomorphology, 81*,330-344.

Drăgut, L., Tiede, D., & Levick, S. R. (2010). ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science, 24*(6), 859-871.

Dronen, S.I. (1984). Windbreaks in the Great Plains. *Northern Journal of Applied Forestry, 3*, 55-59.

Duro, D.C., Franklin, S.E., & Dubé, M.G. (2012). A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5. *Remote Sensing of Environment, 118*, 259-272.

Eby, J.R. (1987). The use of sun incidence angle and infrared reflectance levels in mapping old-growth coniferous forests. In Proceedings, ASPRS-ACSM fall convention: Prospecting new horizons, ASPRS Technical Paper (pp. 36-40). Reno, Nevada, 4-9 October, American Society of Photogrammetric & Remote Sensing, Falls Church Virginia.

Eby, J.R., & Snyder, M.C. (1990). The status of old growth in western Washington: A Landsat perspective. Report, Washington Department of Wildlife, Olympia, Washington.

ESRI Inc. (2009). ArcGIS Desktop: Release 9.3.1. Redlands, CA: Environmental Systems Research Institute.

Falkowski, M.J., Wulder, M.A., White, J.C., & Gillis, M.D. (2009). Supporting large-area, sample-based forest inventories with very high spatial resolution satellite imagery. *Progress in Physical Geography, 33*(3), 403-423.

FAO. (2001). Global forest resources assessment 2000. Main report. Rome: Food and Agriculture Organization of the United Nations. Retrieved from http://www.fao.org/forest-resources-assessment/past-assessments/fra-2000/en/

Fiorella, M., & Ripple, W.J. (1993). Determining successional stage of temperate coniferous forests with Landsat satellite data. *Photogrammetric Engineering and Remote Sensing, 59*, 239-246.

Franklin, S.E. (2001). *Remote Sensing for Sustainable Forest Management*. Boca Raton, FL: CRC Press LLC.

Frederiksen, P., & Lawesson, J.E. (1992). Vegetation types and patterns in Senegal based on multivariate analysis of field and NOAA-AVHRR satellite data. *Journal of Vegetation Science, 3*, 535-544.

Freund, Y., & Schapire, R.E. (1996). Experiments with a new boosting algorithm. In: Machine Learning (pp. 148-156). Proceedings of the Thirteenth International Conference.

Friedl, M.A., & Brodley, C.E. (1997). Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment, 61*, 399-409.

Frohn, R.C. (2006). The use of landscape pattern metrics in remote sensing image classification. *International Journal of Remote Sensing, 27*(10), 2025-2032.

Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N., & Wickham, J. (2011). Completion of the 2006 National Land Cover Database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing, 77*(9), 858-864.

Gaston, G.G., Jackson, P.L., Vinson, T.S., Kolchugina, T.P., Botch, M., & Kobak, K. (1994). Identification of carbon quantifiable regions in the former Soviet Union using unsupervised classification of AVHRR global vegetation index images. *International Journal of Remote Sensing, 15*(16), 3199-3221.

Ghimire, K., Dulin, M.W., Atchison, R.L., Goodin, D.G., & Hutchinson, J.M.S. (2014). Identification of windbreaks in Kansas using object-based image analysis, GIS techniques and field survey. *Agroforestry Systems, 88*, 865-875.

Gislason, P., Benediktsson, J., & J. Sveinsson. (2006). Random Forests for landcover classification. *Pattern Recognition Letters, 27*, 294-300.

Gopal, S., & Woodcock, C.E. (1994). Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing, 60(2)*, 181-188.

Guo, Q., Brandle, J.M., Schoeneberger, M., & D. Buettner. (2004). Simulating the dynamics of linear forests in Great Plains agroecosystems under changing climates. *Canadian Journal of Forest Research, 34*, 2564-2572.

Guo, Q., Kelly, M., Gong, P., & Liu, D. (2007). An object-based classification approach in mapping tree mortality using high spatial resolution imagery. *GIScience & Remote Sensing, 44*(1), 1-24.

Ha, T.V., Amichev, B.Y., Belcher, K.W., Bentham, M.J., Kulshreshtha, S.N., Laroque, C.P., & Van Rees, K.C.J. (2019). Shelterbelt agroforestry systems inventory and removal analyzed by object-based classification of satellite data in Saskatchewan, Canada. *Canadian Journal of Remote Sensing, 45*, 246-263.

Hansen, M.C., DeFries, R.S., Townshend, J.R.G., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach*. International Journal of Remote Sensing, 21*(6-7), 1331-1364.

Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Sohlberg, R., Dimiceli, C., & Carroll, M. (2002). Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. *Remote Sensing of Environment, 83*(1), 303-319.

Hansen, M.H. (1985). Line intersect sampling of wooded strips. *Forest Science, 31*(2), 282-288.

Hansen, M.H. personal communication, April 17, 2013.

Haralick, R.M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-3* (6), 610-621.

Haralick, R.M., & Shapiro, L.G. (1985). Image segmentation techniques. *Computer Vision, Graphics, and Image Processing, 29*(1), 100-132.

Hart, R.H., & Hart, J.A. (1997). Rangelands of the Great Plains before European Settlement. *Rangelands, 19* (1), 4-11.

Hartfield, K.A., Landau, K.I., & van Leeuwen, W.J.D. (2011). Fusion of high resolution aerial mutlispectral and LiDAR data: Land cover in the context of urban mosquito habitat. *Remote Sensing, 3*, 2364-2383.

Hartong, A.L., & Moessner, K.E. (1956). Wooded strips in Iowa. Forest Survey Release 21. USDA Central States Forest Experiment Station, Columbus, OH.

Hay, G., Blaschke, T., Marceau, D., & Bouchard, A. (2003). A comparison of three image-object methods for the multiscale analysis of landscape structure. *ISPRS Journal of Photogrammetry and Remote Sensing, 57*, 327-345.

Hay, G.J., & Castilla, G. (2006). Object-based image analysis: strengths, weaknesses, opportunities and threats (SWOT). International Archives of Photogrammetry, *Remote Sensing and Spatial Information Sciences, 36*, 4.

Hay, G.J., Castilla, G., Wulder, M.A., & Ruiz, J.R. (2005). An automated object-based approach for the multiscale image segmentation of forest scenes. *International Journal of Applied Earth Observation and Geoinformation, 7*, 339-359.

Hay, G.J., Marceau, D.J., Dube, P., & Bouchard, A. (2001). A multiscale framework for landscape analysis: Object-specific analysis and upscaling. *Landscape Ecology, 16*, 471-490.

Hay, G.J., Niemann, K.O., & McLean, G. (1996). An object-specific image-texture analysis of H-resolution forest imagery. *Remote Sensing of Environment, 55*, 108-122.

Hayes, M.M., Miller, S.N., & Murphy, M.A. (2014). High-resolution landcover classification using Random Forest. *Remote Sensing Letters, 5*, 112-121.

Hellesen, T., & Matikainen, L. (2013). An object-based approach for mapping shrub and tree cover on grassland habitats by use of LiDAR and CIR orthoimages. *Remote Sensing, 5*, 558-583.

Hepner, G.F., Logan, T., Ritter, N., & Bryant, N. (1990). Artificial neural network classification using a minimal training set: Comparison to conventional supervised classification. *Photogrammetric Engineering and Remote Sensing, 56*(4), 469-473.

Herold, M., Liu, X., & Clarke, K. (2003). Spatial metrics and image texture for mapping urban land use. *Photogrammetric Engineering and Remote Sensing, 69*(9), 991-1002.

Homer, C., Dewitz, J., Fry, J., Coan, M., Hossain, N., Larson, C., Herold, N., McKerrow, A., VanDriel, J.N., & Wickham, J. (2007). Completion of the 2001 National Land Cover Database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing, 73*(4), 337-341.

Hulet, A., Roundy, B.A., Petersen, S.L., Bunting, S.C., Jensen, R.R., & Roundy, D.B. (2014). Utilizing National Agriculture Imagery Program data to estimate tree cover and biomass of piñon and juniper woodlands. *Rangeland Ecology and Management, 67*, 563-572.

Hurt, R. (1996). Forestry on the Great Plains, 1902-1942. Retrieved from http://www-personal.ksu.edu/~jsherow/hurt2.htm

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Networks, 13*, 411-430.

Iowa DNR. (2017). High resolution land cover of Iowa in 2009. Retrieved from https://geodata.iowa.gov/dataset/high-resolution-land-cover-iowa-2009

Isaacson, D.L., Leckenby, D.A., & Alexander, C.J. (1982). The use of large-scale aerial photography for interpreting Landsat digital data in an elk habitat-analysis project. *Journal of Applied Photogrammetric Engineering, 8*, 51-57.

Johansen, K., Coops, N., Gergel, S., & Stange, Y. (2007). Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification. *Remote Sensing of Environment, 110*, 29-44.

Jones, H.G., & Vaughan, R.A. (2010). *Remote sensing of vegetation: principles, techniques, and applications* (1st ed.), New York, USA: Oxford University Press Inc.

Jose, S. (2009). Agroforestry for ecosystem services and environmental benefits: an overview. Agroforestry Systems, 76, 1-10.

Kauth, R.J., & G.S. Thomas. (1976). The tasseled cap – A graphic description of the spectral-temporal development of agricultural crops as seen by LANDSAT. In Proceedings of the LARS 1976 Symposium on Machine Processing of Remotely Sensed Data, Purdue University.

Kellerman, T.A., Meneguzzo, D.M., Vaitkus, M., White, M., Ossell, R., Sorsen, N., Stannard, J., Gift, T., Cox, J., & Liknes, G.C. (2019). High-resolution land cover of Nebraska (2014). doi:10.2737/RDS-2019-0038

Kettig, R.L., & Landgrebe, D.A. (1976.) Classification of multispectral image data by extraction and classification of homogeneous objects. *IEEE Transactions on Geoscience Electronics, GE-14*(1), 19-26.

Kim, M., Warner, T.A., Madden, M., Atkinson, D.S. (2011). Multi-scale GEOBIA with very high spatial resolution digital aerial imagery: scale, texture and image objects. *International Journal of Remote Sensing, 32*, 2825-2850.

King, R.B. (2002). Land cover mapping: a return to interpretation fundamentals. *International Journal of Remote Sensing, 23*(18), 3525-3545.

Kauth, R., & Richardson, W. (1977). Procedure B: A multisegment training selection and proportion procedure for processing Landsat agricultural data, ERIM Final Report 122700-31-F, Contract NAS 9-14988, Ann Arbor, MI, 158 pp.

Kleinn, C. (2000). On large-area inventory and assessment of trees outside forests. Retrieved from http://www.fao.org/tempref/docrep/fao/x3989e/x3989e02.pdf

Klimm, L.E. (1958). Description of a land use map of Pennsylvania. Tech. Rept. 2.

Knight, J.F., Tolcser, B.P., Corcoran, J.M., Rampi, L.P. (2013). The effects of data selection and thematic detail on the accuracy of high spatial resolution wetland classifications. *Photogrammetric Engineering and Remote Sensing*, 79, 613-623.

Kort, J., & Turnock, R. (1999). Carbon reservoir and biomass in Canadian prairie shelterbelts. *Agroforestry Systems, 44*, 175-186.

Laliberte, A. S., Fredrickson, E. L., & Rango, A. (2007). Combining decision trees with hierarchical object-oriented image analysis for mapping arid rangelands. *Photogrammetric Engineering and Remote Sensing, 73*(2), 197-207.

Laliberte, A.S., Browning, D.M., & Rango, A. (2012). A comparison of three feature selection methods for object-based classification of sub-decimeter resolution UltraCam-L imagery. *International Journal of Applied Earth Observation and Geoinformation, 15*, 70-78.

Land Cover Data Project. Retrieved from https://chesapeakeconservancy.org/conservation-innovation-center-2/high-resolution-data/land-cover-data-project/

Landgrebe, D. A. (1980). The development of a spectral-spatial classifier for earth observational data, *Pattern Recognition, 12*, 165-175.

Landgrebe, D. (1997). The evolution of Landsat data analysis. *Photogrammetric Engineering and Remote Sensing, 63*(7), 859-867.

Lang, S., & Blaschke, T. (2006). Bridging remote sensing and GIS – what are the main supportive pillars? *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences vol.XXXVI-4/C42*, CD-ROM.

Lang, S., & Langanke, T. (2006). Object-based mapping and object-relationship modeling for land use classes and habitats. *Photogrammetrie, Fernerkundung, Geoinformation 10*(1), 5-18.

Lang, S., Schöpfer, E., & Langanke, T. (2009). Combined object-based classification and manual interpretation – synergies for a quantitative assessment of parcels and biotopes. *Geocarto International, 24*(2), 99-114.

Li, X., Myint, S.W., Zhang, Y., Galletti, C., Zhang, X., & Turner II, B.L. (2014). Object-based land-cover classification for metropolitan Phoenix, Arizona, using aerial photography. *International Journal of Applied Earth Observation and Geoinformation, 33*, 321-330.

Li, X., & Shao, G. (2013). Object-based urban vegetation mapping with high-resolution aerial photography as a single data source. *International Journal of Remote Sensing, 34*(3), 771-789.

Li, X., & Shao, G. (2014). Object-based land-cover mapping with high resolution aerial photography at a county scale in Midwestern USA. *Remote Sensing, 6*, 11372-11390.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News. 2/3*, 18-22.

Liknes, G.C., Meneguzzo, D.M., & Kellerman, T.A. (2017). Shape indexes for semi-automated detection of windbreaks in thematic tree cover maps from the central United States. *International Journal of Applied Earth Observation and Geoinformation, 59*, 167-174.

Lillesand, T.M., Kiefer, R.W., & Chipman, J.W. (Eds.). (2008). *Remote Sensing and Image Interpretation* (6th ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Liknes, G.C., Perry, C.H., & Meneguzzo, D.M. (2010). Assessing tree cover in agricultural landscapes using high-resolution aerial imagery. *Journal of Terrestrial Observation, 2*(1), Article 5.

Lister, A., Scott, C., & Rasmussen, S. (2009). Inventory of trees in nonforest areas in the Great Plains states. In M. McWilliams, G. Moisen, & Czaplewski, R. (Comps.), Forest Inventory and Analysis (FIA) Symposium 2008 (pp. 17:1-7).

Lister, A.J., Scott C.T., & Rasmussen S. (2012). Inventory methods for trees in nonforest areas in the Great Plains states. *Environmental Monitoring and Assessment 184*(4):2465-2474.

Liu, Y., Guo, Q., & Kelly, M. (2008). A framework of region-based spatial relations for non-overlapping features and its application in object based image analysis. *ISPRS Journal of Photogrammetry and Remote Sensing, 63*, 461-475.

Loveland, T.R. (2012). History of land-cover mapping. In C.P. Giri (Ed.), *Remote sensing of Land Use and Land Cover: Principles and Applications* (pp. 13-22). Boca Raton, FL: CRC Press, Taylor & Francis Group, LLC.

Loveland, T.R., Merchant, J.W., Reed, B.C,. Brown, J.F., & Ohlen, D.O. (1995). Seasonal land cover regions of the United States. *Annals of the Association of American Geographers, 85*(2), 339-355.

Loveland, T.R., Zhu, Z., Ohlen, D.O., Brown, J.F., Reed, B.C., & Yang, L. (1999). An analysis of the global land cover characterization process. *Photogrammetric Engineering and Remote Sensing, 65*(9), 1021-1032.

Lunetta, R. S. (2004). *Remote Sensing and GIS Accuracy Assessment*. United Kingdom: CRC Press.

Ma, L., Li, M., Ma, X., Cheng, L., Du, P., & Liu, Y. (2017). A review of supervised object-based land-cover image classification. *ISPRS Journal of Photogrammetry and Remote Sensing, 130*, 277–293.

MacFadden, C.H. (1949). Some preliminary notes on the use of the light airplane and the 35 mm camera in geographic field research. *Annals of the Association of American Geographers, 39*(3), 188-200.

Marceau, D.J., Howarth, P.J., Dubois, J.M., & Gratton, D.J. (1990). Evaluation of the grey-level co-occurrence matrix method for land-cover classification using SPOT imagery. *IEEE Transactions on Geoscience and Remote Sensing, 28*(4), 513-519.

Marschner, F.J. (1958). Land use and its patterns in the United States. U.S. Department of Agriculture, Agriculture Handbook, 153, 277 p.

Martínez, A.M., & Kak, A.C. (2001). PCA versus LDA. *IEEE Transactions on Pattern Analysis, 23* (2), 228-233.

Maxwell, A.E., Warner, T.A., Vanderbilt, B.C., & Ramezan, C.A. (2017). Land cover classification and feature extraction from national agriculture imagery program (NAIP) Orthoimagery: A Review. *Photogrammetric Engineering and Remote Sensing, 83*, 737–747.

Maxwell, A.E., Strager, M.P., Warner, T.A., Ramezan, C.A., Morgan, A.N., & Pauley, C.E. (2019). Large-area, high spatial resolution land cover mapping using random forests, GEOBIA, and NAIP orthophotography: Findings and recommendations. *Remote Sensing, 11*, 1409.

Maxwell, A.E., Strager, M.P., Warner, T.A., Zégre, N.P., & Yuill, C.B. (2014). Comparison of NAIP orthophotography and RapidEye satellite imagery for mapping of mining and mine reclamation. *GIScience Remote Sensing, 51,* 301-320.

McGarigal, K., & Marks, B.J. (1995). FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure. USDA Forest Service General Technical Report PNW-351.

Meneguzzo, D.M., Butler, B.J., Crocker, S.J., Haugen, D.E., Moser, W.K., Perry, C.H., Wislon, B.T., Woodall, C.W. (2008). Nebraska's forests, 2005. Retrieved from https://www.nrs.fs.fed.us/pubs/rb/rb_nrs27.pdf

Meneguzzo, D.M., Liknes, G.C., Nelson, M.D. (2013). Mapping trees outside forests using high-resolution aerial imagery: a comparison of pixel- and object based classification approaches. *Environmental Monitoring and Assessment, 185*: 6261-6275.

Meneguzzo, D.M., Lister, A.J., & Sullivan, C. (2018). Summary of findings from the Great Plains Tree and Forest Invasives Initiative. doi: 10.2737/NRS-GTR-177.

Montagnini, F., & Nair, P.K.R. (2004). Carbon sequestration: An underexploited environmental benefit of agroforestry systems. *Agroforestry Systems, 61*, 281-295.

Mora, B., Wulder, M.A., & White, J.C. (2010). Segment-constrained regression tree estimation of forest stand height from very high spatial resolution panchromatic imagery over a boreal environment. *Remote Sensing of Environment, 114*(11), 2474-2484.

Morrison, P.H., Klopfer, D., Leversee, D.A., Socha, C.M., & Ferber, D.L. (1991). *Ancient forests in the Pacific Northwest. Analysis and maps of twelve national forests.* Washington D.C.: The Wilderness Society.

Moskal, L.M., Styers, D.M., & Halabisky, M. (2011). Monitoring urban tree cover using object-based image analysis and public domain remotely sensed data*. Remote Sensing, 3*, 2243-2262.

Muller, E. (1997). Mapping riparian vegetation along rivers: Old concepts and new methods. *Aquatic Botany, 58,* 411.437.

Myeong, S., Nowak, D.J., Hopkins, P.F., & Brock, R.H. (2003). Urban cover mapping using digital, high-spatial resolution aerial imagery. *Urban Ecosystems, 5*, 243-246.

Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment, 115*(5), 1145-1161.

Narumalani, S., Zhou, Y., & Jelinkski, D.E. (1998). Utilizing geometric attributes of spatial information to improve digital image classification. *Remote Sensing Review, 16*, 233-253.

Nebraska Forest Service. (2010). Nebraska statewide forest resource assessment & strategy. Retrieved from http://www.nfs.unl.edu/assessmentstrategy/Nebraska%20SFRAS%20June% 2018%20201 0.pdf

Nelson, M.D., Brewer, M., Meneguzzo, D.M., Clark, K. (2016). Forests of Iowa, 2015. Retrieved from https://www.fs.fed.us/nrs/pubs/ru/ru_fs77.pdf

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., & Wulder, M.A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment, 148*, 42-57.

Olson, C.E., (1960). Elements of photographic interpretation common to several sensors. *Photogrammetric Engineering, 26* (4), 651–656.

O'Neil-Dunne, J.P.M., MacFaden, S., & Pelletier, K. (2011). Incorporating contextual information into object-based image analysis workflows. In ASPRS 2011 Annual Conference Proceedings. Milwaukee, WI.

O'Neil-Dunne, J., MacFaden, S., Royar, A., Reis, M., Dubayah, R., & Swatantran, A. (2014). An object-based approach to statewide land cover mapping. In Proceedings of ASPRS 2014 Annual Conference, Louisville, KY, USA.

O'Neil-Dunne, J., Pelletier, K., MacFaden, S., Troy, A., & Grove, J.M. (2009). Object-based high-resolution land-cover mapping: operational considerations. In 2009 17th International Conference on Geoinformatics (pp. 1-6). IEEE.

O'Neil-Dunne, J., MacFaden, S., & Royar, A. (2014). A versatile, production-oriented approach to high-resolution tree-canopy mapping in urban and

suburban landscapes using GEOBIA and data fusion. *Remote Sensing, 6*, 12837-12865.

Pal, M. (2005). Random forest classifier for remote sensing classification. International Journal of Remote Sensing, 26(1), 217-222.

Pal, N.R., & Pal, S.K. (1993). A review on image segmentation techniques. *Pattern Recognition, 26*, 1277-1294.

Pankiw, J. (2013). The use of object-based classification of high resolution panchromatic satellite imagery for the inventory of shelterbelts in the province of Saskatchewan, Faculty of Graduate Studies and Research. SK, Canada: University of Regina.

Pankiw, J., & Piwowar, J.M. (2010). Seasonality of imagery: the impact on object-based classification accuracy of shelterbelts. *Prairie Perspectives: Geographical Essays, 13*, 39-48.

Pariona, A. Least Densely Populated U.S. States. (2017). Retrieved from worldatlas.com/articles/least-densely-populated-u-s-states.html

Pasher, J., McGovern, M., & Putinski, V. (2016). Measuring and monitoring linear woody features in agricultural landscapes through earth observation data as an indicator of habitat availability. *International Journal of Applied Earth Observation and Geoinformation, 44*, 113-123.

Paull, D.A., Whitson, J.W., Marcotte, A.L., Liknes, G.C., Meneguzzo, D.M., & Kellerman, T.A. (2017). High-resolution land cover of Kansas (2015). doi:10.2737/RDS-2017-0025.

PBS. (n.d.). Surviving the Dust Bowl: The Drought. In *PBS*. Retrieved from https://www.pbs.org/wgbh/americanexperience/features/dustbowl-drought/

Peña-Barragán, J.M., Ngugi, M.K., Plant, R.E., & Six, J. (2011). Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sensing of Environment, 115*(6), 1301-1316.

Perea Moreno, A.J., & Meroño De Larriva, J.E. (2012). Comparison between new digital image classification methods and traditional methods for land-cover mapping. In C.P. Giri (Ed.), *Remote sensing of Land Use and Land Cover: Principles and Applications* (pp. 137-152). Boca Raton, FL: CRC Press, Taylor & Francis Group, LLC.

Perry, C.H., Woodall, C.W., Liknes, G.C., & Schoeneberger, M.M. (2009). Filling the gap: improving estimates of working tree resources in agricultural landscapes. *Agroforestry Systems, 75*, 91-101.

Persello, C., & L. Bruzzone. (2010). A novel protocol for accuracy assessment in classification of very high resolution images. *IEEE Transactions on Geoscience and Remote Sensing, 48*, 1232-1244.

Peterson, D.L., Egbert, S.L., Price, K.P., & Martinko, E.A. (2004). Identifying historical and recent land-cover changes in Kansas using post-classification change detection techniques. *Transactions of the Kansas Academy of Science, 107*, 105-118.

Piwowar, J.M., Amichev, B.Y., & Van Rees, K.C.J. (2017). The Saskatchewan shelterbelt inventory. *Canadian Journal of Soil Science, 97,* 433-438.

Platt, R.V., & Rapoza, L. (2008). An evaluation of an object-oriented paradigm for land use/land cover classification. *The Professional Geographer, 60*(1), 87-100.

Platt, R.V., & Schoennagel, T. (2009). An object-oriented approach to assessing changes in tree cover in the Colorado Front Range 1938-1999. *Forest Ecology and Management, 258*, 1342-1349.

Prasad, A.M., Iverson, L.R., & Liaw, A. (2006). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems, 9,* 181-199.

Qiu, X.M., Wu, S.-S., & Miao, X. (2014). Incorporating road and parcel data for object-based classification of detailed urban land covers from NAIP images. *GIScience Remote Sensing, 51*, 498-520.

Rabben, E.L. (1960). Fundamentals of photo interpretation. In R.N. Colwell (Ed.), *Manual of Photographic Interpretation* (pp. 99-168). Washington D.C.: ASPRS.

Raile, G.K. (1986). Nebraska's second forest inventory. Resource Bulletin NC-96. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station.

Ramachandran Nair, P.K., Mohan Kumar, B., & Nair, V.D. (2009). Agroforestry as a strategy for carbon sequestration. *Journal of Plant Nutrition and Soil Science, 172* (1), 10-23.

Rawat, J.K. (2003). Training manual on inventory of trees outside forests (TOF). Food and Agriculture Organization of the United Nations, Bangkok.

Read, R. (1958). The Great Plains shelterbelt in 1954: a re-evaluation of field windbreaks planted between 1935 and 1942 and a suggested research program. Great Plains Agriculture Council, No. 16.

Richards, J.A., & Jia, X. (1999). *Remote Sensing Digital Image Analysis: An Introduction* (3rd ed.). Berlin: Springer.

Rietveld, B., & Irwin, K. (1996). Agroforestry in the United States. Retrieved from http://www.unl.edu/nac/agroforestrynotes/an01g01.pdf

Ripple, W.J. (1994). Determining coniferous forest cover and forest fragmentation with NOAA-9 advanced very high resolution radiometer data. *Photogrammetric Engineering and Remote Sensing, 60*, 533-540.

Robertson, G., & Mason, A. (2016). Assessing the sustainability of agricultural and urban forests in the United States. Retrieved from https://www.fs.fed.us/research/publications/FS-1067SustainabilityAgUrb.pdf

Robinson, C., Hou, L., Malkin, K., Soobitsky, R., Czawlytko, J., Dilkina, B., & Jojic, N. (2019). Large scale high-resolution land cover mapping with multi-resolution data. Proceedings of the 2019 Conference on Computer Vision and Pattern Recognition (CVPR 2019).

Robinson, E.B., & Dietz, J.L. (2019). Great Plains. In *Encyclopedia Britannica*. Retrieved from https://www.britannica.com/place/Great-Plains

Ryherd, S., & Woodcock, C. (1996). Combining spectral and texture data in the segmentation of remotely sensed imagery. *Photogrammetric Engineering and Remote Sensing, 62*(2), 181-194.

Sauer, T.J. (2009). The Prairie States Forestry Project as a Model for an Effective Global Climate Change Mitigation Project. In L.R. Kellimore (Ed.), *Handbook of Agroforestry Management Practices* (Chapter 22). New York: Nova Science Publishers, Inc.

Sawe, B. (2017). The 50 US states ranked by population. 2017. Retrieved from worldatlas.com/articles/us-states-by-population.html

Schapire, R.E. (1999). A brief introduction to boosting. In Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (pp. 1401–1406).

Schmidt, T.L., & Wardle, T.D. (1998). The forest resources of Nebraska. Research Paper NC-332. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station.

Schoeneberger, M.M. (2005). Agroforestry: working trees for sequestering carbon on ag lands. In K.N. Brooks, & P.F. Folliot (Eds.), Moving Agroforestry into the Mainstream. Proceedings of the 9th North American Agroforestry Conference. St. Paul, MN.

Schoeneberger, M.M. (2009). Agroforestry: working trees for sequestering carbon on agricultural lands. *Agroforestry Systems, 75*, 27-37.

Schoeneberger, M.M., Bentrup, G., de Gooijer, H., Soolanayakanahally, R., Sauer, T., Brandle, J., Zhou, X., & Current, D. (2012). Branching out: agroforestry as a climate change mitigation and adaptation tool for agriculture. *Journal of Soil and Water Conservation, 67*(5), 128A-136A.

Shah, C.A., Anderson, I., Gao, Z., Hao, S., & Leason, A. (2007a). Towards the development of next generation remote sensing technology – ERDAS IMAGINE incorporates a higher order feature extraction technique based on ICA. In Proceedings of the ASPRS 2007 Annual Conference, Bethesda, MD.

Shah, C.A., Varshney, P.K., & Arora, M.K. (2007b). ICA mixture model algorithm for unsupervised classification of remote sensing imagery. *International Journal of Remote Sensing, 28*(8), 1711-1731.

Steiner, D. (1965). Use of air photographs for interpreting and mapping rural land use in the United States. *Photogrammetria, 20*(2), 65-80.

Stone, R.N., & Bagley, W.T. (1961). The forest resource of Nebraska. Forest Survey Release 4. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station.

St. Peter, J., Hogland, J., Anderson, N., Drake, J., & Medley, P. (2018). Fine resolution probabilistic land cover classification of landscapes in the southeastern United States. *ISPRS International Journal of Geo-information, 7*, 107.

Su, W., Li, J., Chen, Y., Liu, Z., Zhang, J., Low, T.M., Suppiah, I., & Hashim, S.A.M. (2008). Textural and local spatial statistics for the object-oriented classification of urban areas using high resolution imagery. *International Journal of Remote Sensing, 29*(11), 3105-3117.

Tansey, K., Chambers, I., Anstee, A., Denniss, A., & Lamb, A. (2009). Object-oriented classification of very high resolution airborne imagery for the extraction of hedgerows and field margin cover in agricultural areas. *Applied Geography, 29*, 145-157.

"The Establishment of a Forest Shelter Belt." (1934). *Science, 80* (2065), p. 91. doi: 10.1126/science.80.2065.91

Thompson, S. (2002). *Sampling* (2nd ed.). New York: John Wiley and Sons, Inc.

Treeratpituk, P., & Giles, C. (2009). Disambiguating authors in academic publications using random forests. Proceedings of the 9th ACM/IEEE-CS joint conference on Digital Libraries, pp. 39-48.

Tomppo, E., Olsson, H., Stahl, G., Nilsson, M., Hagner, O., & Katila, M. (2008). Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sensing of Environment, 112*, 1982-1999.

Troy, A., Grove, J.M., & O'Neil-Dunne, J. (2012). The relationship between tree canopy and crime rates across an urban-rural gradient in the greater Baltimore region. *Landscape and Urban Planning, 106*(3), 262-270.

Tucker, C.J., & Choudhury, B.J. (1987). Satellite remote sensing of drought conditions. *Remote Sensing of Environment, 23*(2), 243-251.

Tucker, C.J., Townshend, J.R.G., & Goff, T.E. (1985). African land cover classification using satellite data. *Science, 227*, 369-375.

USDA Forest Service. 2010. Forest inventory and analysis nation core field guide. vol.1: field data collection procedures for phase 2 plots, version 5.0 Retrieved from http://www.nrs.fs.fed.us/fia/data-collection/

USDA Forest Service, Forest Inventory and Analysis Program, Forest Inventory EVALIDator web-application Version 1.8.0.01. Retrieved from http://apps.fs.usda.gov/Evalidator/evalidator.jsp

USDA National Agricultural Statistics Service. (2009). 2007 Census of Agriculture County Profile – Steele County,MN Retrieved from http://www.agcensus.usda.gov/Publications/2007/Online_Highlights/County_ Profiles/Minnesota/cp27147.pdf

USDA National Agricultural Statistics Service, Census of Agriculture 2017. Retrieved from ttps://www.nass.usda.gov/Publications/AgCensus/2017/index.php

USDA National Agricultural Statistics Service CropScape - Cropland Data Layer. Retrieved from https://nassgeodata.gmu.edu/CropScape/

van der Werff, H.M.A., & van der Meer, F.D. (2008). Shape-based classification of spectrally identical objects. *ISPRS Journal of Photogrammetry and Remote Sensing, 63* (2), 251–258.

Vannier, C., &Hubert-Moy, L. (2014). Multiscale comparison of remote sensing data for linear woody feature vegetation mapping. *International Journal of Remote Sensing, 35* (21), 7376-7399.doi:10.1080/01431161.2014.968683

Vogelmann, J.E., Sohl, T., & Howard, S.M. (1998). Regional characterization of land cover using multiple sources of data. *Photogrammetric Engineering and Remote Sensing, 64*(1), 45-57.

Walker, J.S., & Blaschke, T. (2008). Object-based land-cover classification for the Phoenix metropolitan area: Optimization vs. transportability. *International Journal of Remote Sensing, 29*(7), 2021-2040.

Walsh, S.J. (1980). Coniferous tree species mapping using Landsat data. *Remote Sensing of Environment 9*, 11-26.

Wang, J., & Chang, C. (2006). Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. *IEEE Transactions on Geoscience and Remote Sensing, 44*(6), 1586-1600.

Weather Atlas: Nebraska, USA - Climate data and average monthly weather. Retrieved from https://www.weather-us.com/en/nebraska-usa-climate

Weather Atlas: Kansas, USA – Climate data and average monthly weather. Retrieved from https://www.weather-us.com/en/kansas-usa-climate

Wickham, J., Homer, C., Vogelmann, J., McKerrow, A., Mueller, R., Herold, N., & Coulston, J. (2014). The Multi-Resolution Land Characteristics (MRLC) Consortium – 20 years of development and integration of USA national land cover data. *Remote Sensing, 6*, 7424-7441.

Wiseman, G., Kort, J., & Walker, D. (2009). Quantification of shelterbelt characteristics using high-resolution imagery. *Agriculture, Ecosystems & Environment, 131*, 111-117.

Witharana, C., & Civco, D.L. (2014). Optimizing multi-resolution segmentation scale using empirical methods: Exploring the sensitivity of the supervised discrepancy measure Euclidean distance 2 (ED2). *ISPRS Journal of Photogrammetry and Remote Sensing, 87*, 108-121.

Woodcock, C.E., & Strahler, A.H. (1987). The factor of scale in remote sensing. *Remote Sensing of Environment, 21*, 311-332.

Wulder, M.A., & Franklin, S.E. (Eds.). (2007). *Understanding Forest Disturbance and Spatial Pattern: Remote Sensing and GIS Approaches*. Boca Raton, FL: CRC Press.

Xian, G., Homer, C., & Fry, J. (2009). Updating the 2001 National Land Cover Database land cover classification to 2006 by using Landsat imagery change detection methods. *Remote Sensing of Environment, 113*(6), 1133-1147.

Yan, G., Mas, J. F., Maathuis, B. H. P., Xiangmin, Z., & Van Dijk, P. M. (2006). Comparison of pixel-based and object-oriented image classification approaches: A case study in a coal fire area, Wuda, Inner Mongolia, China. *International Journal of Remote Sensing, 27*, 4039–4055.

Ye, S., Pontius, R.G., & Rakshit, R. (2018). A review of accuracy assessment for object-based image analysis: From per-pixel to per-polygon approaches. ISPRS Journal of Photogrammetry and Remote Sensing, 141, 137–147.

Yu, Q., Gong, P., Clinton, N., Biging, G., Kelly, M., & Schirokauer, D. (2006). Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogrammetric Engineering and Remote Sensing, 72*(7), 799-811.

Zabala, A., Cea, C., & Pons, X. (2012). Segmentation and thematic classification of color orthophotos over non-compressed and JPEG 2000 compressed images. *International Journal of Applied Earth Observation and Geoinformation, 15*, 92-104.

Zhang, Y. (2001). Texture-integrated classification of urban treed areas in high-resolution color-infrared imagery. *Photogrammetric Engineering and Remote Sensing, 67*, 1359-1365.

Zhou, W., & A. Troy. (2008). An object-oriented approach for analyzing and characterizing urban landscape at the parcel level. *International Journal of Remote Sensing, 29*(11), 3119-3135.

Zhou, Y., & Y. Wang. (2006). Extraction of impervious surface area using orthophotos in Rhode Island. ASPRS Annual Conference, Reno, NV.

Zhou, W., Huang, G., Troy, A., & Cadenasso, M. (2009). Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: a comparison study. *Remote Sensing of Environment, 113*, 1769-1777.

Zhou, W., Troy, A., & Grove, M. (2008). Object-based land cover classification and change analysis in the Baltimore Metropolitan Area using multitemporal high resolution remote sensing data. *Sensors, 8*, 1613-1636.

Zhu, Z., & Evans, D.L. (1994). U.S. forest types and predicted percent forest cover from AVHRR data. *Photogrammetric Engineering and Remote Sensing, 60*(5), 525-531.

Zomer, R.J., Trabucco, A., Coe, R., & Place, F. (2009). Trees on Farms: Analysis of Global Extent and geographical patterns of agroforestry. ICRAF Working Paper. ICRAF, Nairobi.